REMARKS

Claims 1-16 and 19-31 are pending.

Claims 2-16 and 19-31 are withdrawn.

Claim 1 is rejected.

Claims 1, 2, 7, 8, 24, 25, 29 and 30 are amended.

Specification

Applicants have amended to specification to add section headings as requested by the examiner.

Applicants have also added the current address of the depository collection on page 7.

No new matter is added.

Claim Objections.

Claims 1, 2, 7, 8, 24, 25, 29 and 30 are amended to italicize all the Latin microorganism names.

No new matter is added.

35 USC 101

Claim 1 is amended to indicate the hand of the inventor. The term "isolated" is inserted into claim 1 as taught on page 7, line 22 of the specification.

No new matter is added.

Thus the claim now reflects the hand of man and overcomes the 101 rejection.

35 USC 112, second paragraph

Claim 1 is rejected as being indefinite. Specifically the examiner believes "a mutant" of strain NCIMB 41164 to be indefinite as the specification does not provide definitions and/or characterization of the mutant as intended.

Regarding the term "mutant" it is known in the art to produce mutants and one skilled in the art could do this readily, whether these are spontaneous mutants through, for example, culturing in a specific way or whether by using genetic mutating techniques.

The applicants enclose for the examiners convenience a chapter from "Molecular Biology: Understanding the Genetic Revolution" by David P. Clark (Elsevier Academic Press, 2005), pages 333-367 describing mutations.

Deposit

Claim 1 is rejected under 35 USC 112, first paragraph as containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

Deposit

Please find enclosed herewith the deposit receipt and the viability statement for the strain NCIMB 41164 which describe all necessary items.

The material has been accepted for deposit under the Budapest treaty on the international Recognition of the Deposit of Microorganisms for the purpose of Patent Procedure and all restrictions on the availability to the public of the material so deposited will be irrevocably removed upon the granting of a patent.

35 USC 102/103

Claim 1 is rejected under 35 USC 102(b) as anticipated by or, in the alternative, under 35 USC 103(a) as obvious over US 5,089,411, Yamada.

Examiner believes Yamada to disclose an isolated *Rhodococcus rhodochrous* strain J-1. Examiner further states that the J-1 microorganism appears to be identical to the presently claimed strain. In the alternative, even if the claimed microorganism is not identical to the references strain and to its subcultrues with regard to some unidentified characteristics, the differences between that which is disclosed and that which is claimed are so slight that the referenced strain J-1 and its subcultrues are likely inherently to possess the same characteristics of the claimed strain.

Applicants respectfully disagree with the examiner's analysis.

The taxonomic identification using the 16S rRNA technique was carried out at the UK National Collection of Industrial and Marine Bacteria. According to the deposit receipt the analyzed stratin BTR 2496 corresponds to the strain NCIMB 41164.

As described on pages 8 and 9 of the present application, comparison of the sequence with the Microseq™ database identified the best match as *Rhodococcus rhodochrous* with a similarity of 97.46%. This is a genus level match, but was most likely to be a strain of *Rhodococcus rhodochrous.* A further search against the public EMBL database identified the best match to *Rhodococcus rhodochrous* with a similarity of 99.698%.

Thus, there is a difference in the base pair sequence, and the strain NCIMB 41164 is novel with respect to the isolated *Rhodococcus rhodochrous* strain J-1 disclosed in US 5,089,411, Yamada.

Furthermore, applicants show that the cultured microorganism shows very high nitrile hydratase activity. Applicants have found that this high nitrile hydratase activity can be achieved if urea or the urea derivative is not present in any substantial amount in the culture medium at the start of the microorganism growth but is introduced later. Applicants have discovered that this late addition of urea or derivative thereof enables highter nitrile hydratase activity to occur in a shorter period of time than if the urea or the urea derivative had been added at the start of culturing.

Table 1 on page 21 shows that the timing of urea addition makes a big difference in nitrile hydratase activity.

While Yamada discloses the addition of urea and derivatives, Yamada teaches urea addition at the beginning of culturing (col. 4, lines 31-34).

Thus the presently claimed *Rhocococcus rhodochrous* strain NCIMB 41164 is novel with respect to the isolated *Rhodococcus rhodochrous* strain J-1 disclosed in US 5,089,411, because of a difference in the base pair sequence as discussed above.

Furthermore, the *Rhocococcus rhodochrous* strain NCIMB 41164 is unobvious in light of Yamada, because the strain differences cannot be considered "so slight that the referenced strain J-1 and its subcultures are likely inherently to possess the same characteristics of the claimed strain". The increased nitrile hydratase activity of the presently claimed new strain is a significant difference . The

particular strain is especially suitable for the production of (meth)acrylamide from (meth)acylonitrile. This is clear from table 1 on page 21 and also instant example 3.

Thus the presently claimed strain is novel and unobvious in light of Yamada.

Rejoin of claims 2-6 and 24.

As the specific strain is novel and unobvious in light of Yamada, the applicants respectfully request rejoin of claims 2-6 and 24, as each of these claims relate to the new strain NCIMB 41164.

Reconsideration and withdrawal of the rejection of claim 1 is respectfully solicited in light of the remarks and amendments *supra*.

Since there are no other grounds of objection or rejection, passage of this application to issue with claim 1 is earnestly solicited.

Applicants submit that the present application is in condition for allowance. In the event that minor amendments will further prosecution, Applicants request that the examiner contact the undersigned representative.

Respectfully submitted,

Shiela A. Loggins

Ciba Corporation
540 White Plains Road
Tarrytown, New York 10591
(914) 785-2768
SAL\22351R1.doc

Shiela A. Loggins
Agent for Applicants
Reg. No. 56,221

Enclosures:

Petition for one(1) month extension of time, deposit receipt for *Rhodococcus rhodochrous* strain NC41165 and copy of pages 335-367 from"Molecular Biology: Understanding the Genetic Revolution" by David P. Clark (Elsevier Academic Press, 2005) describing mutations.

BUDAPEST TREATY ON THE INTERNATIONAL RECOGNITION OF THE DEPOSIT OF MICROORGANISMS FOR THE PURPOSES OF PATENT PROCEDURE

CIBA Specialty Chemicals Water Treatments Ltd
P O Box 38
Cleckheaton Road
Low Moor
Bradford
BD12 0JZ

NAME AND ADDRESS OF DEPOSITOR

INTERNATIONAL FORM

RECEIPT IN THE CASE OF AN ORIGINAL DEPOSIT
issued pursuant to Rule 7.1 by the
INTERNATIONAL DEPOSITARY AUTHORITY
identified at the bottom of this page

I. IDENTIFICATION OF THE MICROORGANISM

Identification reference given by the DEPOSITOR:

Rhodococcus rhodochrous
BTR 2496

Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY:

NCIMB 41164

II. SCIENTIFIC DESCRIPTION AND/OR PROPOSED TAXONOMIC DESIGNATION

The microorganism identified under I above was accompanied by:

[ ] a scientific description

[X] a proposed taxonomic designation

(Mark with a cross where applicable)

III. RECEIPT AND ACCEPTANCE

This International Depositary Authority accepts the microorganism identified under I above, which was received by it on 5 March 2003 (date of the original deposit)[1]

IV. RECEIPT OF REQUEST FOR CONVERSION

The microorganism identified under I above was received by this International Depositary Authority on (date of the original deposit) and a request to convert the original deposit to a deposit under the Budapest Treaty was received by it on (date of receipt of request for conversion)

V. INTERNATIONAL DEPOSITARY AUTHORITY

Name: NCIMB Ltd.,

Address: 23 St Machar Drive,
Aberdeen,
AB24 3RY,
Scotland.

Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorised official(s):

Terence Dodd

Date: 17 July 2003

[1] Where Rule 6/4(d) applies, such date is the date on which the status of International Depositary Authority was acquired.

Form BP/4 (sole page)

BUDAPEST TREATY ON THE INTERNATIONAL RECOGNITION OF THE DEPOSIT OF MICROORGANISMS FOR THE PURPOSES OF PATENT PROCEDURE

CIBA Specialty Chemicals Water Treatments Ltd
P O Box 38
Cleckheaton Road
Low Moor
Bradford
BD12 0JZ

NAME AND ADDRESS OF THE PARTY TO WHOM THE VIABILITY STATEMENT IS ISSUED

INTERNATIONAL FORM

VIABILITY STATEMENT
issued pursuant to Rule 10.2 by the INTERNATIONAL DEPOSITARY AUTHORITY identified on the following page

I. DEPOSITOR	II. IDENTIFICATION OF THE MICROORGANISM
Name: AS ABOVE Address:	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY: NCIMB 41164 Date of the deposit or of the transfer[1]: 5 March 2003
III. VIABILITY STATEMENT	
The viability of the microorganism identified under II above was tested on 6 March 2003 [2]. On that date, the said microorganism was: [3] [X] viable [ ] no longer viable	

[1] Indicate the date of the original deposit or, where a new deposit or a transfer has been made, the most recent relevant date (date of the new deposit or date of the transfer).

[2] In the cases referred to in Rule 10.2(a)(ii) and (iii), refer to the most recent viability test.

[3] Mark with a cross the applicable box.

IV. CONDITIONS UNDER WHICH THE VIABILITY TEST HAS BEEN PERFORMED[4]
V. INTERNATIONAL DEPOSITARY AUTHORITY
Name: NCIMB Ltd., Address: 23 St Machar Drive, Aberdeen, A24 3RY, Scotland. Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorised official(s): Terence Dundon Date: 17 July 2003

[4] Fill in if the information has been requested and if the results of the test were negative.

by mutation, there is a back-up copy which can be used to produce the correct protein. This often suppresses the potential defect, unless the mutation is dominant. It has been estimated that a typical human carries enough harmful mutations to total approximately eight lethal equivalents per genome. Put another way, if humans were haploid, with only a single copy of each gene, the average person would be dead eight times over. Due to mutations accumulated over the centuries, all humans are genetically different from their ancestors.

The Major Types of Mutation

Many different types of mutation occur. Some affect a single base, others affect large segments of DNA.

A single mutation is a single event and a multiple mutation is the result of several events. A single mutational event, however large or complex its effect, is regarded as a single mutation. A mutation that involves only a single base is known as a **point mutation**. A **null mutation** totally inactivates a gene; the expression "null mutation" is a genotypic term. Complete absence of a gene product may or may not cause a detectable phenotype. A **tight mutation** is one whose phenotype is clear-cut. The complete loss of a particular enzyme may result in no product in a particular biochemical pathway. For example, the complete inability of a bacterium to grow if provided with a certain sugar is an example of a tight mutation. A **leaky mutation** is one where partial activity remains. For example, 10 percent residual enzyme activity might allow a bacterium to still grow, albeit very slowly.

The sequence of a DNA molecule may be altered in many different ways. Such mutations have a variety of outcomes that depend in part on the nature of the change and in part on the role of the DNA sequence that was altered. The major types of sequence alteration are as follows, and will be discussed separately below:

Base substitution: one base is replaced by another base.

Insertion: one or more bases are inserted into the DNA sequence.

Deletion: one or more bases are deleted from the DNA sequence.

Inversion: a segment of DNA is inverted, but remains at the same overall location.

Duplication: a segment of DNA is duplicated; the second copy usually remains at the same location as the original.

Translocation: a segment of DNA is transferred from its original location to another position either on the same DNA molecule or on a different DNA molecule.

Mutations may affect genes that encode proteins, genes encoding non-coding RNA, regulatory sequences and recognition sites.

Much of the discussion below considers what happens when mutations occur within genes that encode proteins. However it is important to realize that mutations may also occur within those genes whose products are tRNA, rRNA, or other non-translated RNA molecules. Alterations in these molecules may have drastic effects on ribosome function, splicing or other vital processes. Furthermore, mutations may also fall within promoter sequences or other regulatory sites on the DNA that do not actually encode any gene product. Nonetheless, such regulatory sites are important for gene expression and altering them may have major effects.

base substitution Mutation in which one base is replaced by another
deletion Mutation in which one or more bases is lost from the DNA sequence
duplication Mutation in which a segment of DNA is duplicated
insertion Mutation in which one or more extra bases are inserted into the DNA sequence
inversion Mutation in which a segment of DNA has its orientation reversed, but remains at the same location
leaky mutation Mutation where partial activity remains
null mutation Mutation that totally inactivates a gene
point mutation Mutation that affects a single base pair
tight mutation Mutation whose phenotype is clear-cut due to the complete loss of function of a particular gene product
translocation Mutation in which a segment of DNA is transferred from its original location to another site on the same or a different DNA molecule

Best Available Copy

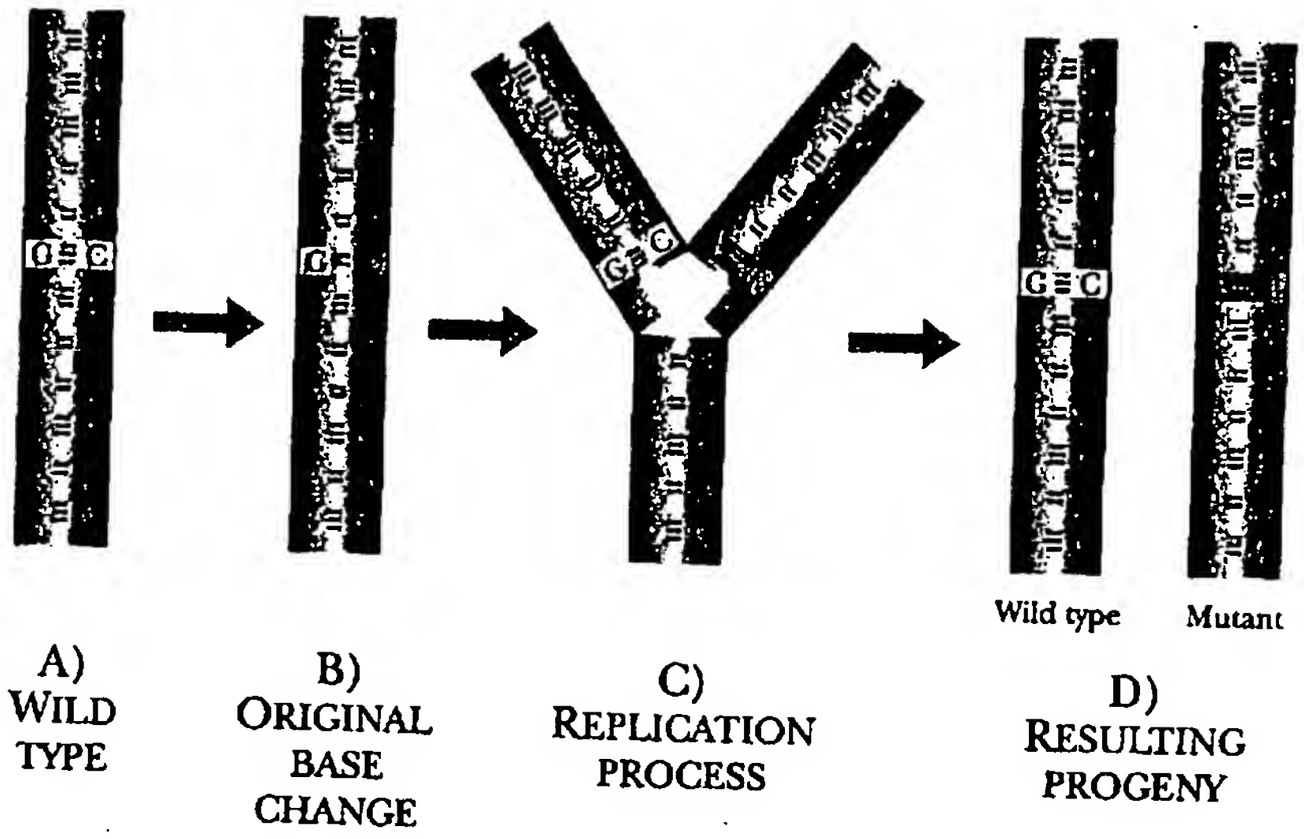


FIGURE 13.01 ***Segregation of Base Alterations in DNA***

A) A mutation has occurred causing a C to be replaced by a T. B) During replication, one DNA molecule (in yellow) matches the original base and the other strand (green) matches the mutated base. C) The strands segregate into the progeny, giving one wild type and one mutant.

Base Substitution Mutations

If one base is replaced by another, a base substitution mutation has occurred. These may be subdivided into **transitions** and **transversions**. In a transition a pyrimidine is replaced by another pyrimidine (i.e., T is replaced by C or vice versa) or a purine is replaced by another purine (i.e., A is replaced by G or vice versa). A transversion occurs when one base is replaced by another of a different type; for example, a pyrimidine is replaced by a purine or vice versa.

DNA molecules are double stranded. If a mutation occurs and a single base is replaced with another, the DNA molecule will temporarily contain a pair of mismatched bases (Fig. 13.01). When the DNA molecule replicates, complementary bases will be incorporated into the new strands opposite the bases making up the mismatch. The result is one wild-type daughter molecule and one mutant DNA molecule.

When mutations are induced by experimental treatment, it is necessary to allow the cells time to divide after treatment before imposing any selection. This allows the original DNA strands to separate and the cell to make new DNA molecules that are either fully wild-type or fully mutant. This process is sometimes referred to as **segregation**, as the originally mutated cell segregates the mutation and the wild-type into separate daughter cells upon cell division.

Missense Mutations May Have Major or Minor Effects

When a change in the base sequence alters a codon so that one amino acid in a protein is replaced with a different amino acid, this is called a **missense mutation**. Overall, this is the most frequent outcome of changing a single base. The severity of a missense mutation depends on the location and the nature of the amino acid that was substituted.

Mutant versions of genes are numbered as described in Ch. 1. Thus *genX123* refers to the 123rd mutation isolated in the gene, *genX*. A mutation which results in a codon for one amino acid being replaced by another may be written *genX123* (Arg185Leu),

missense mutation Mutation in which a single codon is altered so that one amino acid in a protein is replaced with a different amino acid
segregation Replication of a hybrid DNA molecule (whose two strands differ in sequence) to give two separate DNA molecules, each with a different sequence
transition Mutation in which a pyrimidine is replaced by another pyrimidine or a purine is replaced by another purine
transversion Mutation in which a pyrimidine is replaced by a purine or vice versa

Best Available Copy

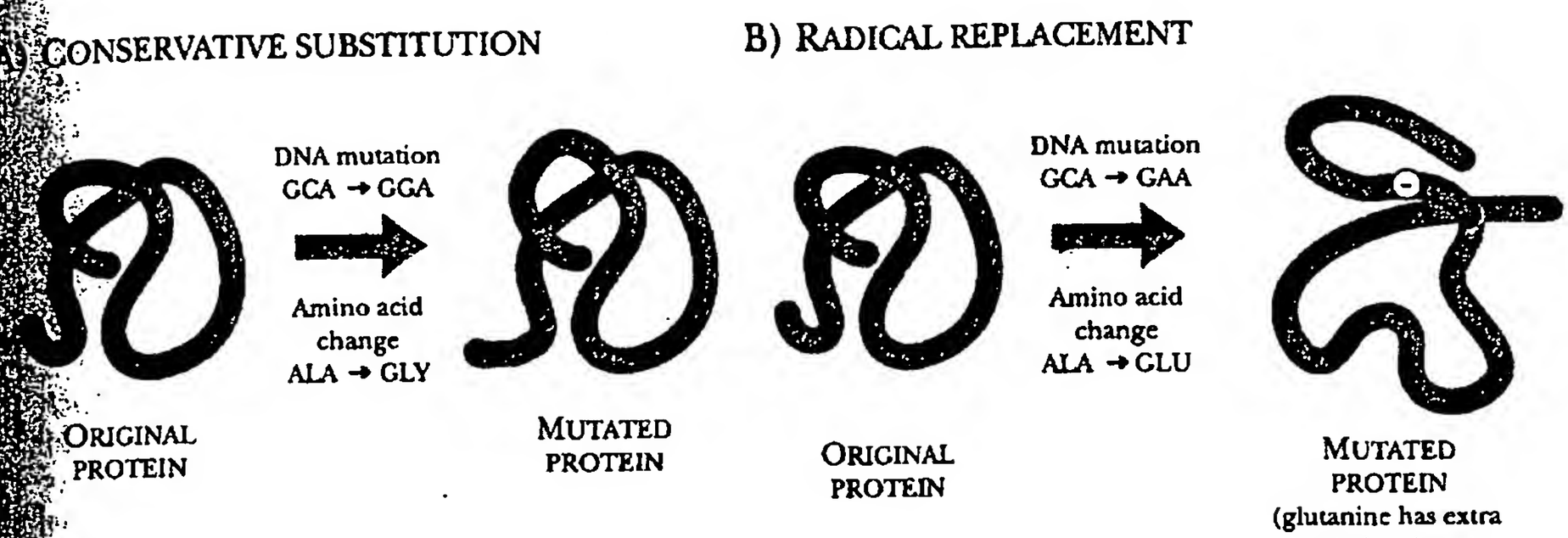


FIGURE 13.02 ***Conservative Substitution and Radical Replacement***

A) A mutation resulting in DNA change from GCA to GGA will result in the conservative substitution of an alanine for a glycine. Since both amino acids have similar properties, it is likely that the mutant protein will fold similarly to the wild type. B) A mutation resulting in the substitution of a glutamate for an alanine is a radical replacement as the glutamate has an extra negative charge that will probably cause the protein to fold quite differently from the wild type.

or in one-letter code (R185L). This indicates that arginine at position 185 has been replaced by leucine.

Proteins must assume their correct three-dimensional structure in order to function properly. Moreover, most proteins, especially enzymes, contain an active site whose role is critical. This region contains relatively few of the many amino acids that make up a typical protein. Sequence comparison of the same protein from different organisms usually shows that only the amino acids in a few positions are invariant or nearly so. These highly conserved amino acid residues generally include those in the active site(s) plus others that are critical for correct folding of the protein. Although the protein must fold up correctly, the precise identity of the amino acids at many positions may vary substantially without causing major changes in overall structure. Thus, mutations that alter active site residues will usually have major effects. Mutations that alter residues important for structure will also have a major impact. However, mutations affecting less vital parts of the protein will often have minor effects and substantial activity may remain.

Replacing an amino acid with a chemically similar one often has little effect on a protein.

The chemical properties of the original amino acid and the one replacing it in the mutant are also important. Suppose the codon UCU, which codes for serine, is changed to ACU, which codes for threonine. Both serine and threonine are small, hydrophilic amino acids with hydroxyl groups. Replacing one amino acid with another that has similar chemical and physical properties is known as a **conservative substitution**. Swapping serine for threonine in the less critical regions of a protein will probably not alter its structure radically and the protein may still work, at least partially. In rare instances, the protein may actually work better. On the other hand, if the exchange is made in a critical region of the protein, such as the active site, even a conservative substitution may completely destroy activity. Nonetheless, since the critical regions of most proteins occupy only a small proportion of the total sequence, most conservative substitutions will be relatively mild and usually non-lethal (Fig. 13.02A).

Replacing an amino acid with one that has very different properties often causes significant damage to the protein.

Replacing one amino acid with another that has different chemical and physical properties is known as a **radical replacement** (Fig. 13.02B). Suppose the codon GUA, which codes for valine, is changed to GAA; this then yields a glutamic acid. This

conservative substitution Replacement of an amino acid with another that has similar chemical and physical properties
radical replacement Replacement of an amino acid with another that has different chemical and physical properties

Best Available Copy

replaces a bulky hydrophobic residue with a smaller, hydrophilic residue that carries a strong negative charge. Under most circumstances, replacing valine with glutamic acid will seriously cripple or totally incapacitate most proteins. If the residue in question is on the surface of the protein, it is sometimes possible to get away with a radical replacement, provided that the change does not affect a critical binding site or alter the solubility of the protein too drastically.

Mutant proteins may sometimes be defective only under certain conditions, such as high temperature.

An interesting and sometimes useful type of missense mutation is the **temperature sensitive (ts) mutation**. As its name indicates, the mutant protein folds properly at low temperature (the "permissive" temperature) but is unstable at higher temperatures and unfolds. Consequently, the protein is inactive at the higher or "restrictive" temperature. If a protein is essential, a missense mutation will often be lethal to the cell. However, a temperature sensitive mutant can be grown and used for genetic experiments at the lower permissive temperature, where it remains alive. To analyze the damage caused by the mutation, the temperature is then shifted upward to the restrictive temperature at which the protein is inactivated and the organism may eventually die. An example in the fruit fly, *Drosophila*, is the *para(ts)* mutation. This affects a protein that forms sodium channels necessary for transmitting nerve impulses. At high temperatures the mutant protein is inactive and the flies are paralyzed. At lower temperatures, they are capable of normal flight (Fig. 13.03).

Mutations whose effects vary depending on a variety of environmental conditions are well known.

Naturally occurring temperature sensitive mutations have given rise to the patterns of fur coloration in some animals. Many light colored animals have black tips to their paws, tails, ears and noses. This is due to a temperature-sensitive mutation in the enzyme that synthesizes melanin, the black skin pigment of mammals. In these cases, the mutant enzyme is inactive at normal mammalian body temperature, but active at the lower temperatures found at the extremities. Consequently, melanin is made only in the cooler outlying parts of the body (Fig. 13.04).

Mutations whose effects vary depending on the environment are known as **conditional mutations**. Cold-sensitive mutations do occur but are much rarer than high or normal temperature-sensitive mutations. Multi-subunit proteins are often held together by hydrophobic patches on the surfaces of the subunits (see Ch. 7). The hydrophobic interaction is weaker at lower temperatures. It is therefore possible to get altered proteins, whose hydrophobic bonding is weaker, that fail to assemble at low temperatures but are normal at higher temperatures. For example, microtubule proteins are temperature dependent. Microtubules are cylinders made from the helical assembly of the monomer tubulin. In *Saccharomyces cerevisiae*, residues whose mutation caused cold sensitivity were concentrated at the interfaces between adjacent alpha-tubulin subunits. Mutations that respond to the osmotic pressure or ionic strength of the medium are also known.

Nonsense Mutations Cause Premature Polypeptide Chain Termination

Not all codons encode amino acids. Three (UAA, UAG and UGA) are stop codons that signal the end of a polypeptide chain. A **nonsense mutation** occurs when the codon for an amino acid is mutated to give a stop codon. Suppose that the codon UCG for serine is changed by replacing the middle base, C, with A. This gives the stop codon UAG. When the ribosome translates the mRNA, it comes to the mutant codon that used to be serine. But this is now a stop codon, so the ribosome stops and the rest of the protein does not get made. Release factor recognizes the premature stop codon and releases the partly-made polypeptide. Hence, nonsense mutations are sometimes called **chain termination mutations**. Usually, the shortened polypeptide

chain termination mutation Same as nonsense mutation
conditional mutation Mutation whose phenotypic effects depend on environmental conditions such as temperature or pH
nonsense mutation Mutation due to changing the codon for an amino acid to a stop codon
temperature-sensitive (ts) mutation Mutation whose phenotypic effects depend on temperature

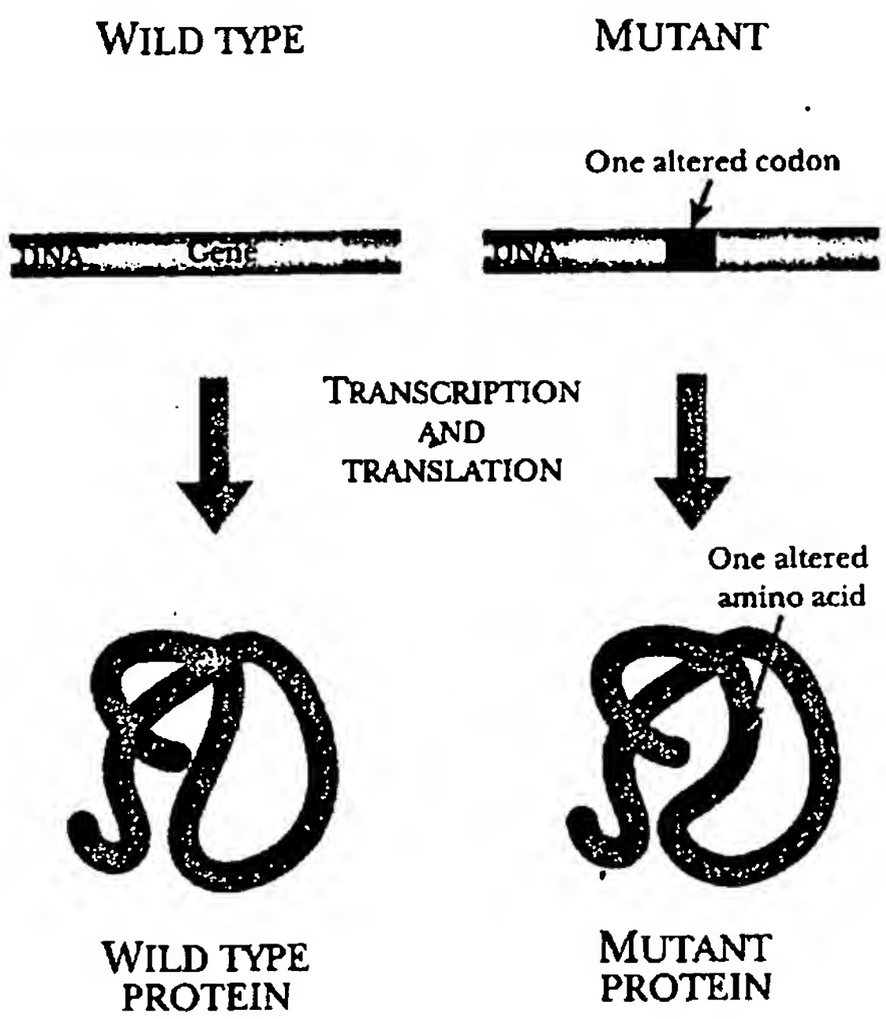


FIGURE 13.03
Temperature Sensitive Mutation

The wild-type gene encodes a protein that folds similarly at high and low temperature. The mutant protein folds normally at low temperature but unfolds at high temperature, and consequently, no longer works properly.

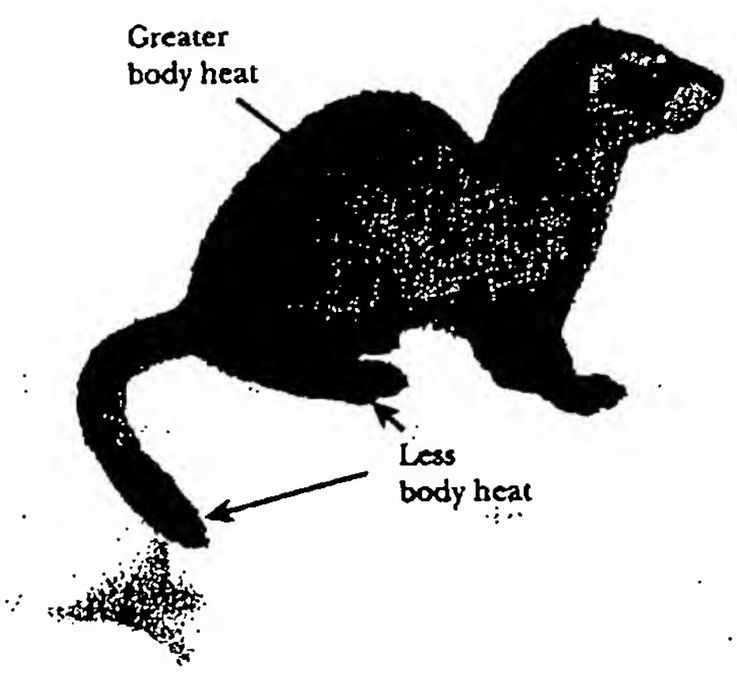


FIGURE 13.04
Temperature-Sensitive Fur Coloration

Some animals change colors in their extremities when body temperature is lowered in response to seasonal temperature change. The mutant gene for melanin is active at lower temperatures.

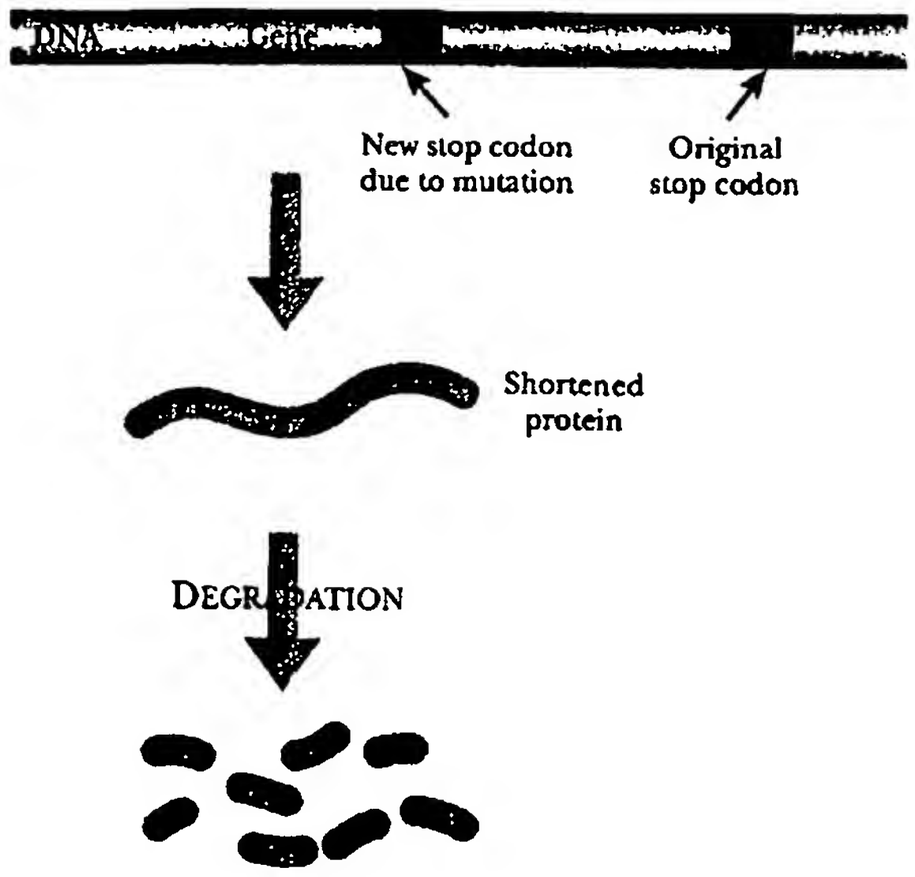


FIGURE 13.05 *Nonsense Mutation*

A mutation in the DNA has produced a new stop codon that causes premature termination of the protein. The shortened protein remains unfolded and is usually detected by the cell and degraded.

chain cannot fold properly (Fig. 13.05). Such misfolded proteins are detected and degraded by the cell (see Ch. 8). The result, in practice, is normally the total absence of this particular protein. Nonsense mutations are often lethal if they affect important proteins.

Deletion Mutations Result in Shortened or Absent Proteins

Mutations that remove one or more bases are known as deletions and those that add extra bases are known as insertions. Clearly, the effect of a deletion (or insertion) depends greatly on how many bases are removed (or inserted). In particular, we should distinguish between point mutations where one (or a very few) bases are affected, and gross deletions and insertions that affect long segments of DNA. Point deletions and insertions may have major effects due to disruption of the reading frame—see below. Here we will consider the effects of larger deletions.

Deletions are indicated by the symbol Δ or by DE. Thus Δ(*argF-lacZ*) or DE(*argF-lacZ*) indicates a deletion of the region (of the *E. coli* chromosome in this case) from the *argF* to the *lacZ* gene. Obviously, deletion of the DNA sequence for a whole gene means that no mRNA and no protein will be made (Fig. 13.06). If the protein is essential, then the deletion will be lethal. Large deletions may remove part of a gene, an entire gene or several genes. Deletions may also remove part or all of the regulatory region for a gene. Depending on the precise region removed, gene expression may be decreased or increased. For example, a deletion that removes the binding site for a repressor may result in a large increase in activity of the gene in question. Thus loss of DNA may result in elevated activity. Again, deletions may remove largely functionless DNA, such as the non-coding sequences between genes or the introns found within genes. In this case, the effects may be small or marginal.

Deletions may remove critical segments of DNA or largely functionless regions of DNA.

Deletion mutations are surprisingly frequent. About 5 percent of spontaneous mutations in bacteria such as *E. coli* are deletions. Although bacteria lack introns and the intergenic spacer regions are very short, it is still possible to generate non-lethal deletions of considerable size. The main reason is that many genes are needed only under certain limited environmental conditions. Thus deletions of the entire *lac* operon in *E. coli* prevent the organism from using lactose as a source of carbon, yet have no other deleterious effects.

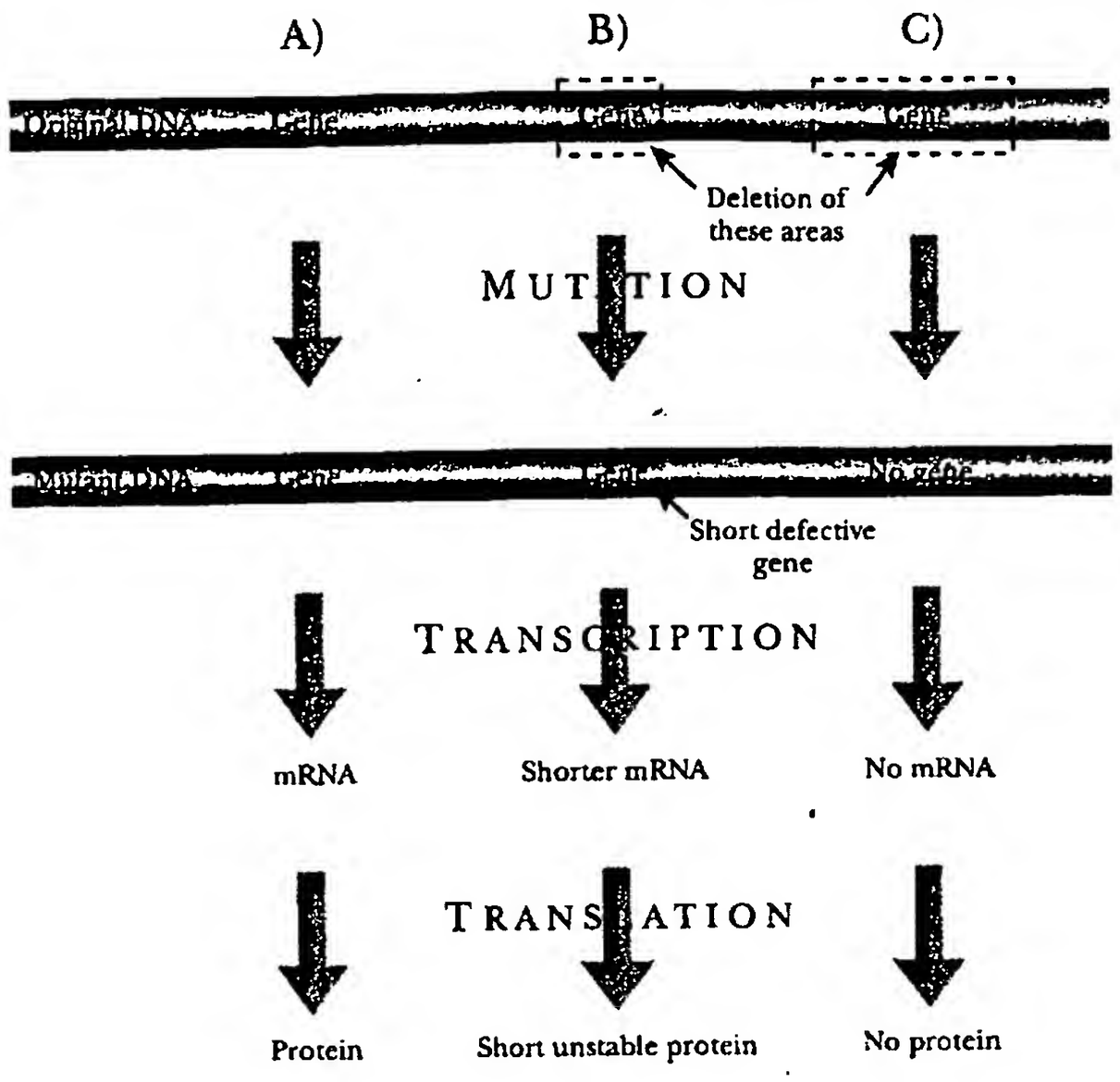


FIGURE 13.06 ***Effects of Deletion Mutations***

A) The wild-type gene produces a normal mRNA and a normal protein. B) A large deletion causes a shorter mRNA and a short unstable protein. C) Deletion of an entire gene results in no mRNA and no protein.

Insertion Mutations Commonly Disrupt Existing Genes

Most naturally occurring insertions are due to mobile elements, including transposons, retroposons and certain viruses.

Genes may also be inactivated by insertions of DNA. If a foreign segment of DNA is inserted into the coding region, then the gene is said to be disrupted. Usually the gene will be completely inactivated; however, the precise result will vary depending on the length and sequence of the inserted DNA (Fig. 13.07) as well as on its precise location. Thus, if an insertion occurs very close to the 3′-end of a gene, most of the coding sequence will remain intact and sometimes a more or less functional protein may still be made. The cause of insertion mutations may be divided into two distinct categories. Some of these mutations are the result of **mobile genetic elements**, usually thousands of bases long, inserting themselves into a gene. Other insertion mutations, usually only one or a few bases long, are caused by mutagenic chemicals or by mistakes made by DNA polymerase. These short insertions are discussed below under "Frameshift Mutations."

Large insertions within operons often prevent transcription of genes downstream from the insertion site.

Under natural conditions, most insertion mutations are due to mobile genetic elements inserting themselves into the DNA. Such elements include insertion sequences, **transposons** and **retroposons** (see Ch. 15) and certain viruses that may integrate their DNA into the host chromosome (see Ch. 17). Insertions are indicated by the symbol :: between the target gene and the inserted element. Thus *lacZ*::Tn10 indicates the insertion of the transposon Tn10 into the *lacZ* gene. Insertion of such large genetic elements disrupts the target gene and completely disrupts its proper function. The presence of transposons greatly increases the frequency of various other DNA rearrangements, such as deletions and inversions. The precise mechanisms are uncertain but are probably due to abortive transposition attempts (see Ch. 15).

Transposable elements and viruses usually contain multiple transcriptional terminators. Consequently, RNA polymerase cannot transcribe through them. Therefore

mobile genetic element A discrete segment of DNA that is able to change its location within larger DNA molecules by transposition or integration and excision
retrotransposon or retroposon A transposable element that uses reverse transcriptase to convert the RNA form of its genome to a DNA copy
transposon Same as transposable element, although the term is usually restricted to DNA-based elements that do not use reverse transcriptase

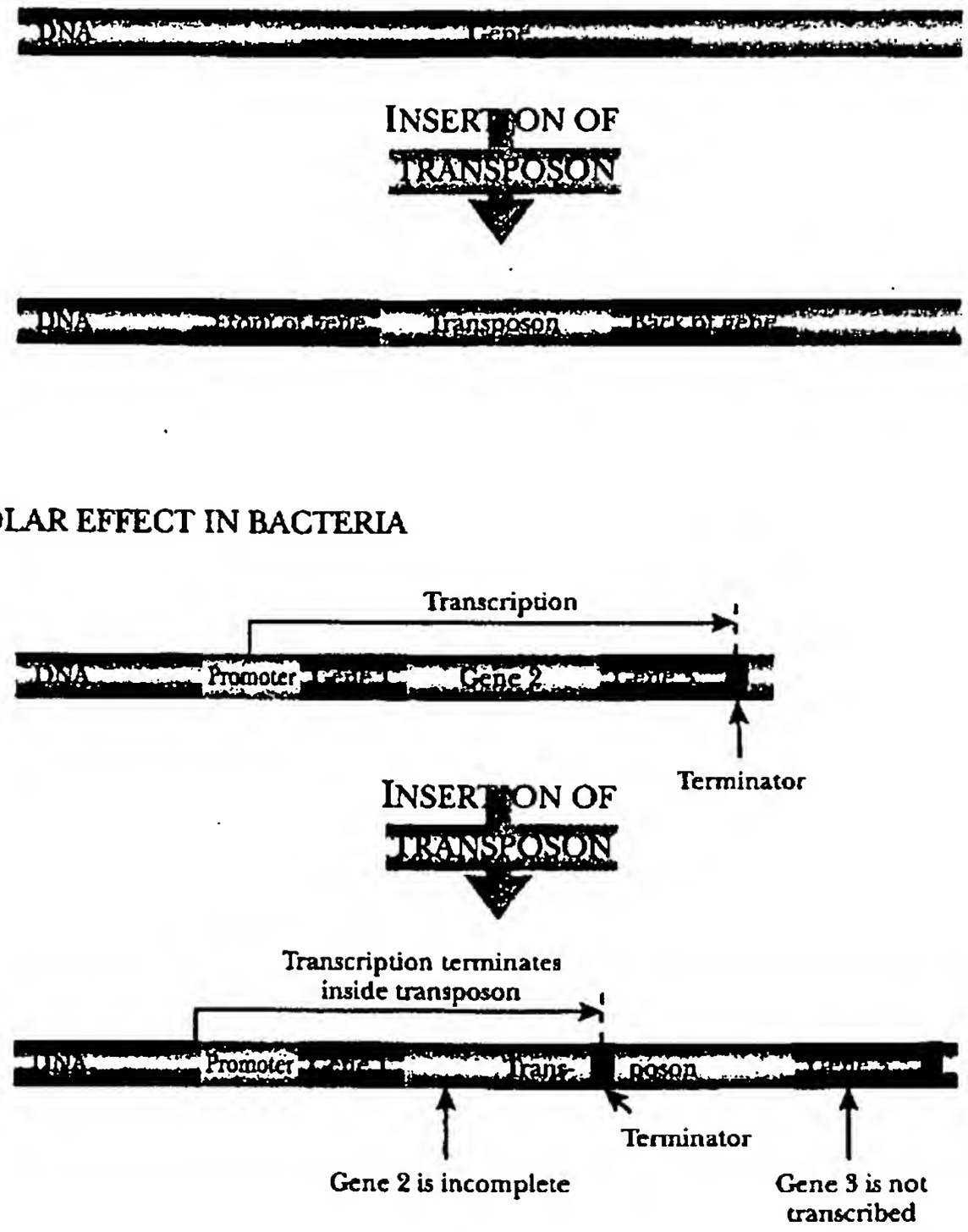


FIGURE 13.07 ***Effects of Insertion Mutations***

A) Insertion of a transposon into the middle of a gene interrupts the coding sequence. B) Insertion of a transposon into the second gene of a bacterial operon with three genes. Gene 1 is the only gene correctly transcribed since the transposon disrupts gene 2 and causes premature termination. Gene 3 will not be transcribed, although its coding sequence is still intact.

their presence blocks transcription of any other downstream genes that share the same promoter as the gene that suffered the insertion event. This effect is referred to as **polarity**. Since bacterial genes are often found clustered in operons and are co-transcribed onto the same mRNA (see Ch. 6), they are much more likely than eukaryotic genes to show polarity effects due to insertions.

Occasionally, insertions may activate genes. If an insertion occurs in the recognition site for a repressor, binding of the repressor will be prevented and activation of the gene may result. In addition, a few transposons are known to have promoters close to their ends, facing outwards (Fig. 13.08). Insertion of these may activate a previously silent gene. Examples are known of "cryptic" genes that have potentially functional gene products that cannot be expressed due to defective promoters. Thus the *bgl* operon of *E. coli* is inactive in the wild type and only expressed in mutants when a transposon carrying an outward-facing promoter is inserted just upstream of the operon and reactivates it.

polarity When the insertion of a segment of DNA affects the expression of downstream genes, usually by preventing their transcription

Best Available Copy

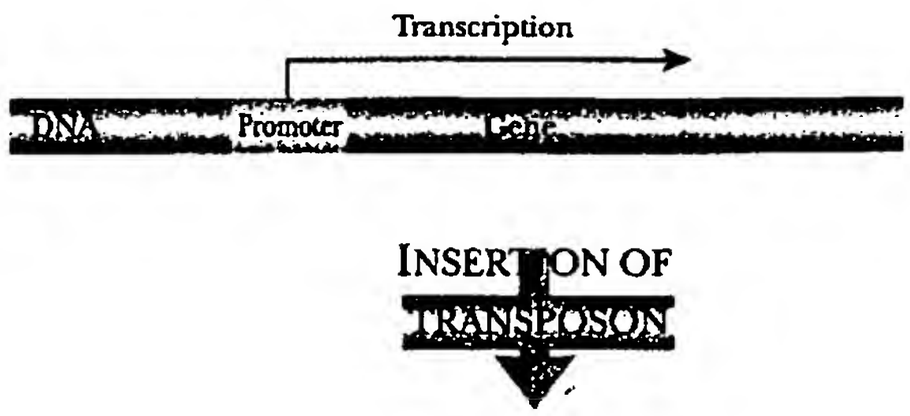


FIGURE 13.08 ***Unusual Activating Effects of Insertion Mutations***

A) The gene shown is under the control of its own promoter. B) A transposon is inserted between the normal promoter and the structural gene. The gene is now expressed under control of a promoter carried by the transposon.

Frameshift Mutations Sometimes Produce Abnormal Proteins

Bases are read as codons, that is in groups of three, when translated into amino acids during protein synthesis (Ch. 8). The introduction or removal of one or two bases can have drastic effects since the alteration changes the reading frame of the afflicted gene. If a single base of a coding sequence is inserted or removed, the reading frame for all codons following the insertion or deletion (Fig. 13.09) will be changed. The result will be a completely garbled protein sequence. Such **frameshift mutations** usually completely destroy the function of a protein, unless they occur extremely close to the far end. Insertion or deletion of two bases also changes the reading frame and alters protein function.

A mutation that alters the reading frame usually disrupts protein function completely.

However, insertion or deletion of three bases adds or removes a whole codon and the reading frame is retained. Apart from the single amino acid that is gained or lost, the rest of the protein is unchanged. If the deleted (or inserted) amino acid is in a relatively less vital region of the protein, a functional protein may be made. Adding or deleting more than three bases will give a similar result as long as the number is a multiple of three. In other words, a whole number of codons must be added or subtracted to avoid the consequences of changing the reading frame.

DNA Rearrangements Include Inversions, Translocations, and Duplications

An inversion is just what its name implies, an inverted segment of DNA (Fig. 13.10A). Reading a stretch of DNA backwards results in drastic changes. Inversions within genes are usually highly detrimental. On the other hand, inversions do not always disrupt genes. If the endpoints of an inversion are in intergenic DNA, then inversion of a DNA segment carrying one or more intact genes, together with their promoters, may have only mild effects. In this case the orientation of the gene(s) will be reversed relative to the rest of the chromosome and it will be transcribed in the opposite direction.

frameshift Mutation in which the reading frame of a structural gene is altered by insertion or deletion of one or a few bases

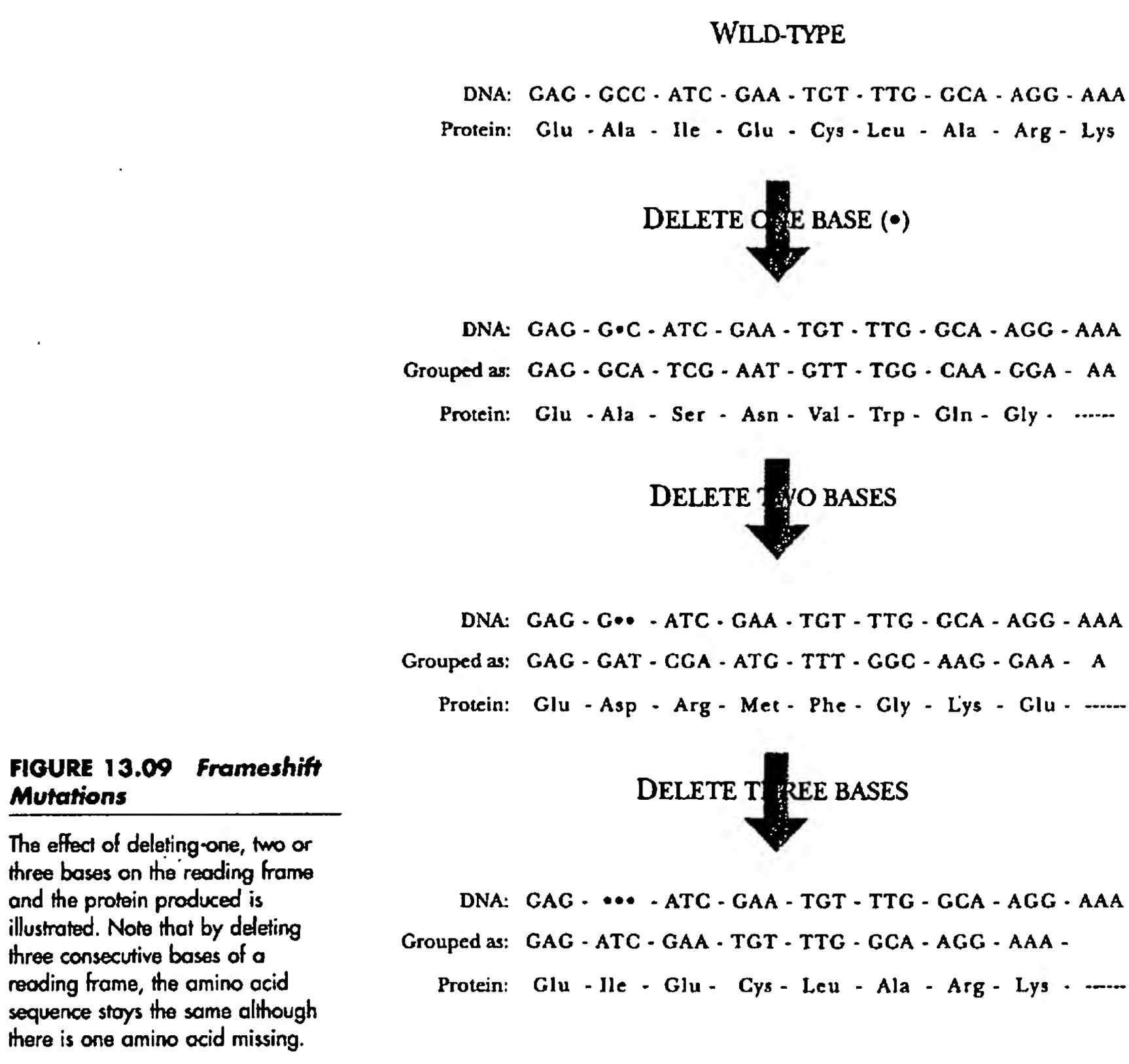


FIGURE 13.09 ***Frameshift Mutations***

The effect of deleting one, two or three bases on the reading frame and the protein produced is illustrated. Note that by deleting three consecutive bases of a reading frame, the amino acid sequence stays the same although there is one amino acid missing.

A translocation is the removal of a section of DNA from its original position and its insertion in another location, either on the same chromosome or on a completely different chromosome (Fig. 13.10B). If an intact gene is merely moved from one place to another, it may still work and little damage may result. On the other hand, if, for example, half of one gene is moved and inserted into the middle of another gene, the results will be doubly chaotic.

A duplication occurs when a segment of DNA is duplicated and both copies are retained. In most cases, the duplicate is located just following the original copy (Fig. 13.10C); in other words, there is a **tandem duplication**. Duplication within a gene will seriously disrupt the gene product, whereas duplication of a large segment of DNA may generate extra copies of genes. Such duplication followed by sequence divergence is thought to be a major source of new genes over the course of evolution (see Ch. 20).

tandem duplication Mutation in which a segment of DNA is duplicated and the second copy remains next to the first

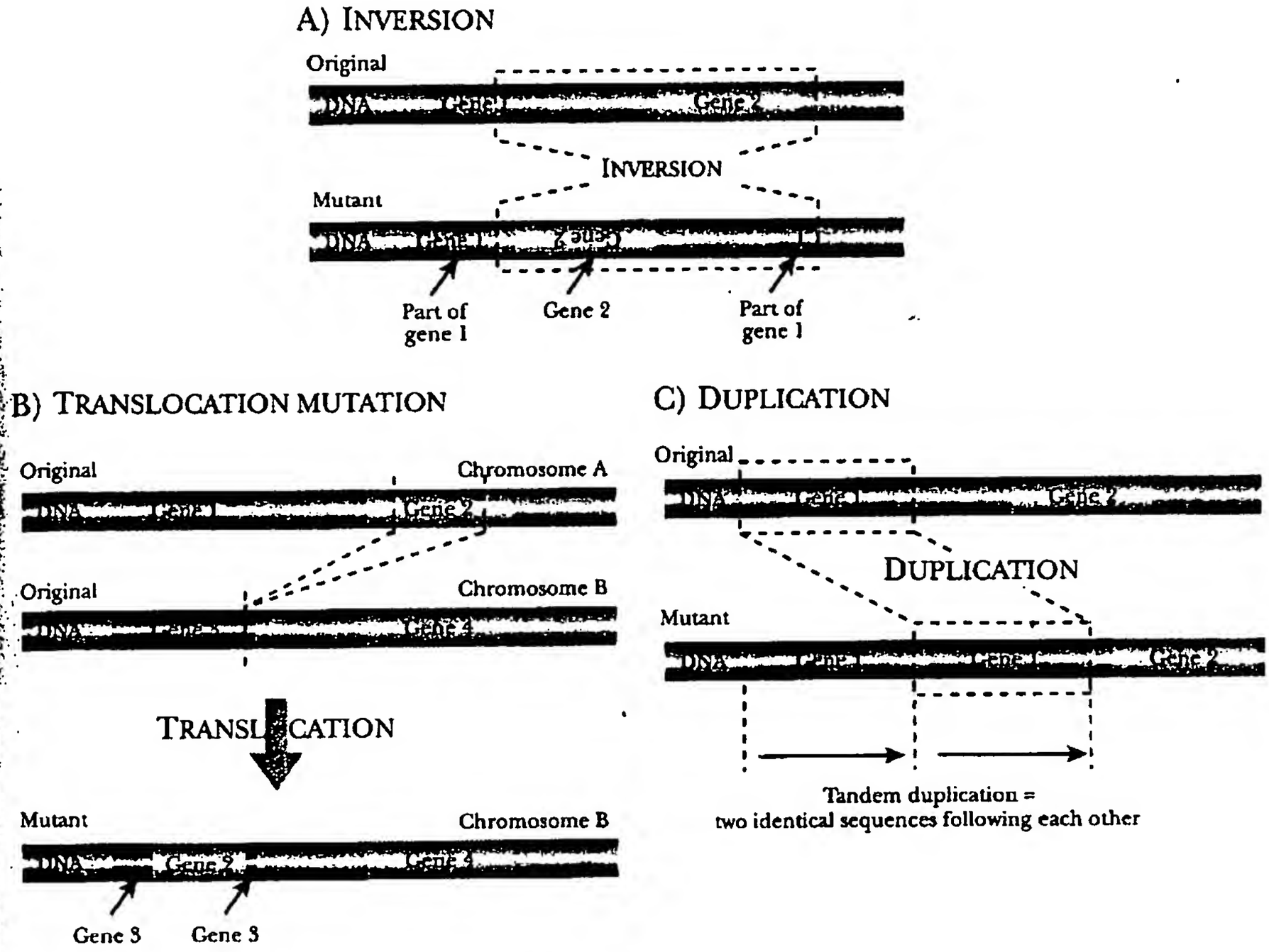


FIGURE 13.10 ***Inversions, Translocations and Duplications***

A) The inversion shown encompasses gene 2 and part of gene 1. The inverted regions are indicated by the backward spelling. B) Two chromosomes are shown with four genes. Part of gene 2 is moved from its location on chromosome A to chromosome B to a position that splits gene 3. C) Gene 1, along with some non-coding DNA, is duplicated in a tandem duplication.

Phase Variation Is Due to Reversible DNA Alterations

A few genes are switched on or off by inverting the regulatory sequences in front of them.

Phase variation is due to the reversible inversion of a DNA segment. When the DNA segment faces forward, a promoter close to its end can transcribe downstream genes (Fig. 13.11). If the segment is inverted, the promoter now faces backwards and the genes cannot be transcribed. Interconversion of the two alternative forms occurs at high frequency. It is catalyzed by a specific enzyme, an **invertase**, that recognizes specific sequences at the two ends of the invertible segment. Thus the system alternates between two phases, one in which the gene(s) is switched off, and the other in which it is turned on. Usually just the regulatory sequences are inverted and only rarely are structural genes actually disrupted by phase variation.

The flip-flop rate for phase variation is about 1 in 10,000 bacterial cells per generation, about a hundred times more frequent than most point mutations in the same organism. Since an inversion is undeniably a change in DNA sequence, which is by definition a mutation, phase variation may be viewed both as a form of genetic regulation and as an aberrant type of mutation.

The best known examples of phase variation occur in bacteria such as *Salmonella* and *E. coli*, where the switch is between expressing two alternative surface proteins.

invertase (Strictly, DNA invertase) An enzyme that recognizes specific sequences at the two ends of an invertible segment and inverts the DNA between them
phase variation Reversible inversion of a segment of DNA leading to differences in gene expression

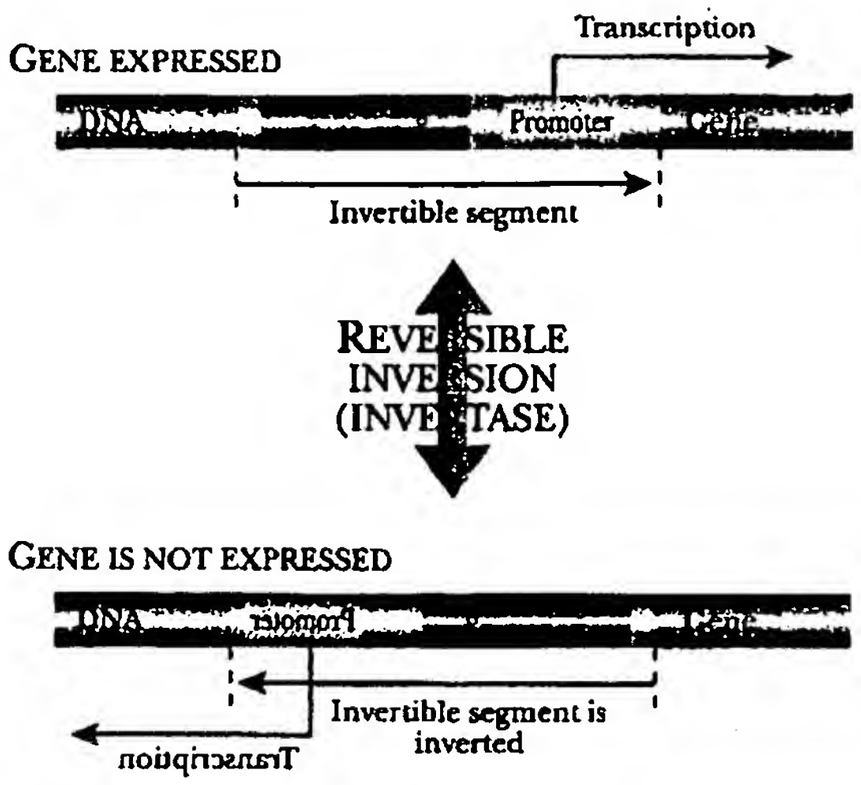


FIGURE 13.11 ***Inversion Causes Phase Variation***

Some segments of DNA can vary in orientation. In the form at the top the DNA segment is oriented so that the promoter is in the correct position to transcribe the gene. In the inverted form (below) the gene is not expressed because the promoter faces the wrong direction.

In *Salmonella*, the flagella can be made from two alternative proteins, H1 and H2. The genes for H2 protein and the rh1 repressor lie downstream of the promoter on the invertible segment (Fig. 13.12). When these are expressed, rh1 represses the gene for H1 protein, which is located elsewhere on the bacterial chromosome. When the segment inverts, due to Hin invertase, the H2 and rh1 genes are no longer expressed. Lack of the rh1 repressor allows instead expression of the H1 gene. Thus either H1 or H2 is expressed, but never both at once. These flagellar proteins (known historically as H-antigens) are strong antigens and found on the surface of bacteria. Periodically switching at random between different flagellar proteins helps bacteria avoid detection by the mammalian immune system.

Pathogenic strains of *E. coli* often bind to intestinal cells by thin helical filaments of protein known as pili or fimbriae. These are also strong surface antigens and are often subject to phase variation by a similar mechanism as described above. Appropriately enough, certain viruses that infect enteric bacteria also use phase variation to change the tail fiber proteins that recognize bacterial surface proteins during infection. This phase variation results in different host ranges for the alternative types of virus particle. For example, bacteriophage Mu flip-flops between recognizing *E. coli* and the closely related *Citrobacter*.

Silent Mutations Do Not Alter the Phenotype

Many mutations have no effect on the phenotype—they are "silent".

A **silent mutation** is an alteration in the DNA sequence that has no effect on the operation of the cell and is therefore not so much silent as invisible from the outside. In other words, silent mutations do not alter the phenotype. One obvious kind of silent mutation is a base change occurring in the non-coding DNA between genes. Therefore, no genes are damaged and no proteins are altered. Higher organisms possess intervening sequences, the introns, within many of their genes. Since introns are cut out and discarded when the messenger RNA is made, most alterations to the sequence of an intron will not affect the final protein.

Not all base changes in an intron are harmless. Changes in the few critical bases at the splice recognition sites will result in failure to splice out the intron or in aberrant splicing. This will give a severely damaged protein product when the misspliced mRNA is translated. Further, many of the small nucleolar RNAs are derived from introns in other genes (Ch. 12). In addition, occasional cases are known where an intron

silent mutation An alteration in the DNA sequence that has no effect on the phenotype

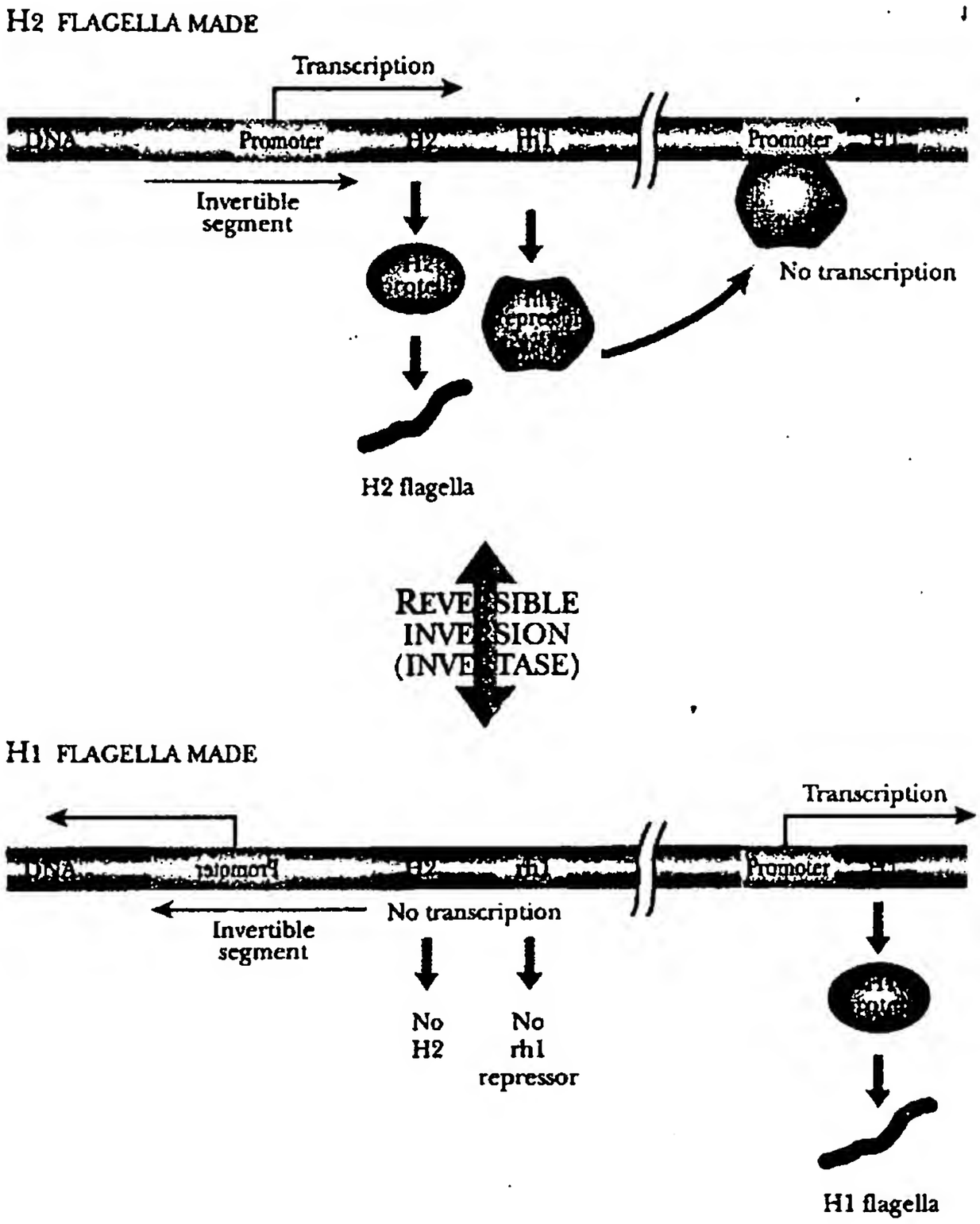


FIGURE 13.12 ***Phase Variation of Flagella in Salmonella***

The top diagram illustrates the state where flagellar protein H2 is made. The invertible segment has its promotor aligned to express H2 and the rh1 repressor, which then represses the gene for H1. Therefore in this orientation only H2 is made. In the bottom diagram the invertible segment is reversed. Neither H2 nor the repressor for H1 are made. Thus H1 is free to be expressed. The two forms of the gene are interconvertable.

is needed to guide base modifications to a neighboring exon (see Ch. 12). Nevertheless, most base changes within most introns are silent mutations.

The third main type of silent mutation occurs within the coding region of a gene and does get passed on to the messenger RNA. Remember that each codon, or group of three bases, is translated into a single amino acid in the final protein product. However, because there are 64 different codons, most of the 20 possible amino acids have more than one codon (see Codon Table, Fig. 8.02). So a base change that converts the original codon into another codon that codes for the same amino acid will have no effect on the final structure of the protein.

Since many amino acids have several codons, changing the third base of a codon often leaves the amino acid unchanged.

For example, the amino acid alanine has four codons: GCU, GCC, GCA and GCG. (Note that the sequences are discussed in RNA language; these are the codons as found on mRNA.) Since they all have GC as the first two bases, any codon of the form GCN (N = any base) will give alanine. A mutation in an original GCC sequence changing the last C to an A or a G or a U results in change in the sequence of the codon, but there is no change in the amino acid produced (alanine) in the resulting protein. Many other amino acids (such as valine, threonine and glycine) also have sets of four codons in which the last base does not matter. This pattern is referred to as **third base redundancy**. Very often, altering the third base of a codon has no effect on the protein that will be made. In other words, about a third of single base substitutions will be silent,

third base redundancy Situation where a set of four codons all code for the same amino acid and thus the identity of the third codon base makes no difference during translation

even if they occur within the protein coding region of a gene. [Note: the term "codon degeneracy" refers to the fact that a single amino acid may be encoded by multiple codons. In most, but not all cases, this is due to third base redundancy. However, both arginine and serine have six codons each; these inevitably differ among themselves by more than just the third base.]

Third base mutations that do not alter protein identity can sometimes have effects due to differential codon usage and tRNA bias. Different codons for the same amino acid often vary in their frequency of use and the corresponding tRNAs are often present in levels that are related to the frequency of codon usage. Consequently, if changing the third base converts a frequently used codon to a rarely used codon for the same amino acid, translation may be slowed due to shortage of the appropriate tRNA. (Conversely, changing a rare to a common codon may speed up translation.) This effect is usually only significant for proteins that are made at such high levels that their synthesis uses a significant proportion of the available tRNA in the cell.

Chemical Mutagens Damage DNA

DNA may be damaged by a variety of chemicals and by radiation.

Mutations that are caused by agents that damage the DNA are known as **induced mutations**. Agents that mutate DNA are called **mutagens** and are of three main types: mutagenic chemicals, radiation and heat. Even if there are no dangerous chemicals or radiation around, mutations still occur, though less frequently. These are **spontaneous mutations**. Some of these are due to errors in DNA replication. The enzymes of DNA replication are not perfect and sometimes make mistakes. In addition, DNA undergoes certain spontaneous chemical reactions (alterations) at a low but detectable rate and this rate goes up with increasing temperature.

The most common mutagens are toxic chemicals that react with DNA and alter the chemical structure of the bases. For example, EMS (ethyl methane sulfonate) is widely used by molecular biologists to mutagenize growing cells. It adds an ethyl group to bases in DNA and so changes their shape and their base-pairing properties. Nitrite converts amino groups to hydroxyl groups and so converts the base cytosine to uracil (Fig. 13.13). Nitrite is used experimentally to mutate purified DNA, such as a cloned gene carried on a plasmid, while the plasmid is in the test tube. The mutagenized DNA is then transferred back into a cell to identify the mutations that were generated. During DNA replication, the DNA polymerase misidentifies these altered bases and puts in the wrong bases in the new complementary strand of DNA it is making (Fig. 13.13).

Base analogs are mistaken by the cell for the natural nucleic acid bases.

Base analogs are chemical mutagens that mimic the bases found in natural DNA. For example, bromouracil resembles thymine in shape. It is converted by the cell to the DNA precursor, bromouridine triphosphate, which DNA polymerase inserts where thymine should go. Unfortunately, bromouracil can flip-flop between two alternative shapes (Fig. 13.14). In its alternate form, bromouracil resembles cytosine and pairs with guanine. If bromouracil is in its misleading form when DNA polymerase arrives, a G will be put into the new strand opposite the bromouracil instead of A.

Some mutagens imitate the structure of a base pair rather than a single base. For example, **acridine orange** has three rings and is about the size and shape of a base pair. Acridine orange is not chemically incorporated into the DNA. Instead, it squeezes in between the base pairs in the DNA (Fig. 13.15), a process referred to as **intercalation**. During DNA replication, the DNA polymerase mistakes the intercalating agent for a base pair and puts in an extra base when making the new strand. As discussed above, insertion of an extra base will change the reading frame of the protein encoded by a

Intercalating agents result in the insertion of an extra base pair during DNA replication.

acridine orange A mutagenic agent that acts by intercalation
base analog Chemical mutagen that mimics a DNA base
induced mutation Mutation caused by external agents such as mutagenic chemicals or radiation
intercalation Insertion of a flat chemical molecule between the bases of DNA, often leading to mutagenesis
mutagen Any agent, including chemicals and radiation, that can cause mutations
spontaneous mutation Mutation that occurs "naturally" without the help of mutagenic chemicals or radiation

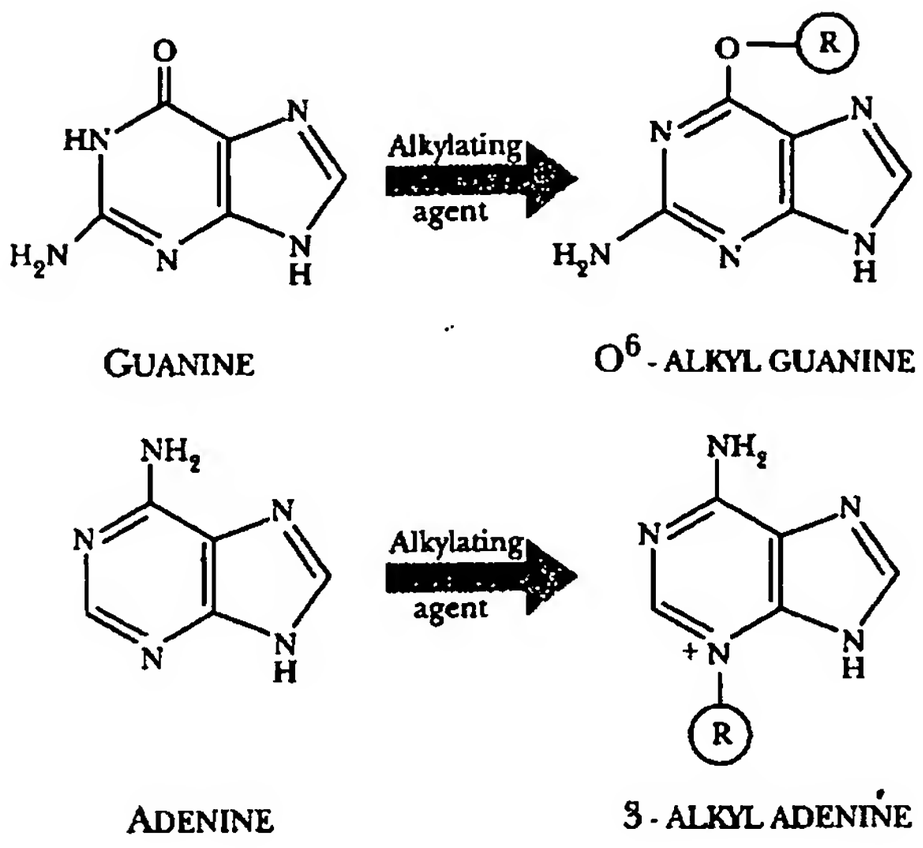


FIGURE 13.13 ***Base Alteration by Chemical Mutagens***

A) Alkylating agents alter the structure of bases by adding alkyl groups. B) Nitrite will convert cytosine to uracil (which pairs with adenine).

FIGURE 13.14 ***Bromouracil Acts as a Base Analog***

Bromouracil has two alternative forms, one of which (left) looks like thymine and pairs with adenine; the other (right) looks like cytosine and pairs with guanine.

gene. Since this will completely destroy the function of the protein, intercalating agents are highly hazardous mutagens.

A **teratogen** is an agent that causes abnormal development of the embryo, which results in gross structural defects. Teratogens may or may not cause mutations. The most famous example is thalidomide, which resulted in the birth of malformed

teratogen An agent that causes abnormal embryo development leading to gross structural defects or monstrosities

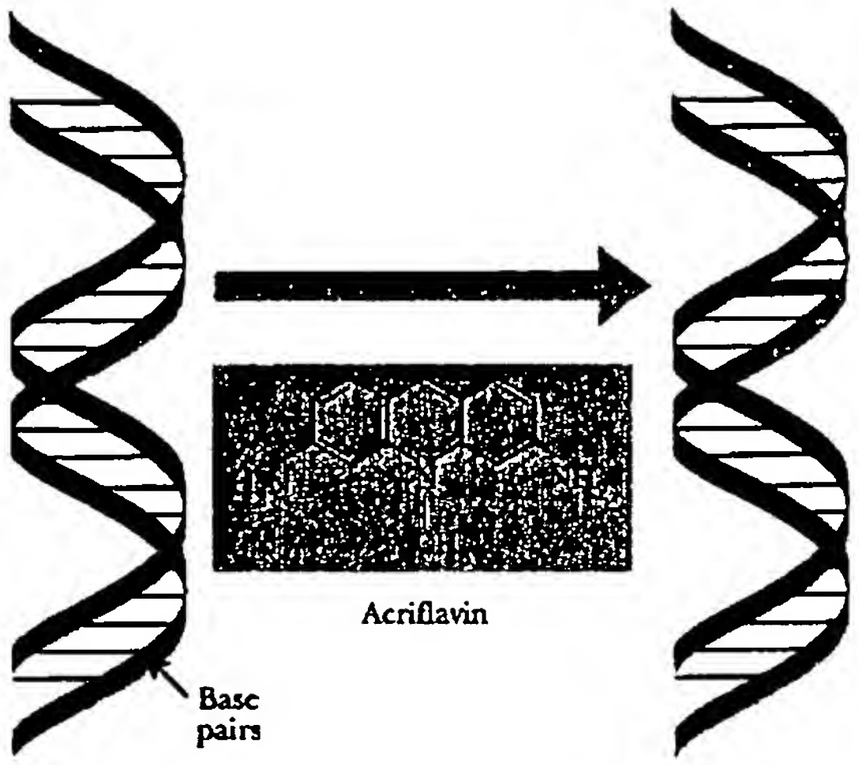


FIGURE 13.15
Intercalating Agents

An intercalating agent, such as acriflavin, can insert itself between base pairs and mimic a whole extra base pair. When replication occurs, the intercalating agent causes an extra base pair to be inserted into the new DNA. Commercial acriflavin is actually a mixture of the structure shown plus the derivative without the N-methyl group.

children with missing limbs. Thalidomide interferes with the development of embryos as opposed to causing mutations. Although the mechanism responsible for the malformations remains uncertain, it is known that thalidomide prevents blood vessels from forming (i.e. it is anti-angiogenic), which may partly explain the drug's ability to cause birth defects.

Radiation Causes Mutations

High energy radiation damages DNA.

Some types of radiation cause mutations. High frequency electromagnetic radiation, ultraviolet radiation (UV light), X-rays and gamma rays (γ-rays), directly damage DNA. X-rays and γ-rays are **ionizing radiation**; that is, they react with water and other molecules to generate ions and free radicals, notably hydroxyl radicals. Ionizing radiation is responsible for about 70 percent of the radiation damage to DNA. The other 30 percent of the radiation damage is due to direct interaction of X-rays and γ-rays with DNA itself. In the early days of molecular biology, X-rays were often used to generate mutations in the laboratory. X-rays tend to produce multiple mutations and often yield rearrangements of the DNA, such as deletions, inversions and translocations.

Ultraviolet radiation promotes formation of thymine dimers.

Ultraviolet radiation is electromagnetic radiation with wavelengths from 100 to 400 nm. It is nonionizing and acts directly on the DNA. The bases of DNA show an absorption peak at around 254 nm and UV close to this wavelength is absorbed very efficiently by DNA. In particular, UV causes two neighboring pyrimidine bases to cross-react with each other to give dimers. Thymine dimers are especially frequent (Fig. 13.16). Although DNA polymerase can proceed by skipping over thymine dimers, this leaves a single-stranded region that needs repairing. The repair process in turn causes the insertion of incorrect bases in the newly synthesized strand (see Ch. 14 for details on error-prone repair). This therefore results in mutation.

Ultraviolet radiation is emitted by the sun. Most of it is absorbed by the ozone layer in the upper atmosphere, so it does not reach the surface of the earth. Damage to the ozone layer by the chlorinated hydrocarbons used in aerosol sprays and refrigerants has allowed more UV radiation to reach the surface of this planet, especially in certain areas. This has probably contributed to the increased frequency of skin cancer noted in recent years.

In addition to electromagnetic radiation, there are other forms of radiation, such as the α-particles and β-particles emitted by radioactive materials along with

ionizing radiation Radiation that ionizes molecules that it strikes

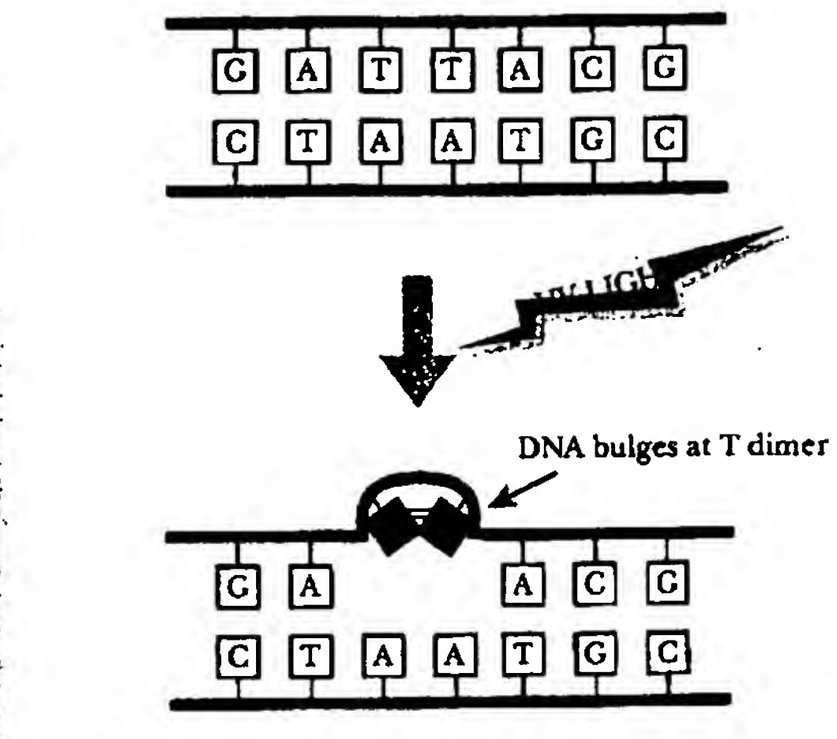


FIGURE 13.16 ***Thymine Dimers***

A) Ultraviolet light (UV) sometimes results in the formation of a thymine dimer (red). B) The detailed chemical structure of the thymine dimer is shown.

γ-rays. Most α-particles are too weak even to penetrate skin but β-particles may cause significant damage to DNA and other biological molecules. However, α-emitters can be mutagenic if they have entered the body, for example by being breathed in or swallowed.

Spontaneous Mutations Can Be Caused by DNA Polymerase Errors

DNA polymerase makes spontaneous mistakes that result in mutations.

The enzymes that replicate DNA during cell division are not perfect. They make errors at a rate that is low, but nonetheless significant over a long period of time. As discussed in Ch. 5, DNA polymerases carry out **proofreading** and check recently inserted nucleotides for mistakes before moving on. In some cases, the proofreading ability is part of the polymerase itself. In other cases, it is due to an accessory protein such as the DnaQ protein associated with *E. coli* DNA polymerase III. Cells carrying mutations that abolish or damage these proofreading abilities show much higher rates of spontaneous mutation. Genes that give rise to altered mutation rates when they themselves are mutated are known as **mutator genes**. Hence, *E. coli dnaQ* mutants were originally named *mutD* (for mutator D).

The error rate for DNA replication in *Escherichia coli* is approximately one base in 10 million. About 20 times as many errors occur in the lagging strand as in the leading strand. This probably results from DNA polymerase I having a less effective proofreading capability than DNA polymerase III. The lagging strand is made discontinuously (see Ch. 5) and the gaps are filled in by DNA polymerase I, whereas the leading strand is all made by PolIII.

DNA polymerase may slip when replicating short sequence repeats.

In addition to putting in an occasional wrong base, DNA polymerase may very rarely omit bases or insert extra bases. This is due to strand slippage. If a run of several identical bases occurs, the template strand and newly synthesized strand of DNA may

mutator gene Gene whose mutation alters the mutation frequency of the organism, usually because it codes for a protein involved in DNA synthesis or repair

proofreading Process that checks whether the correct nucleotide has been inserted into new DNA. Usually refers to DNA polymerase checking whether it has inserted the correct base

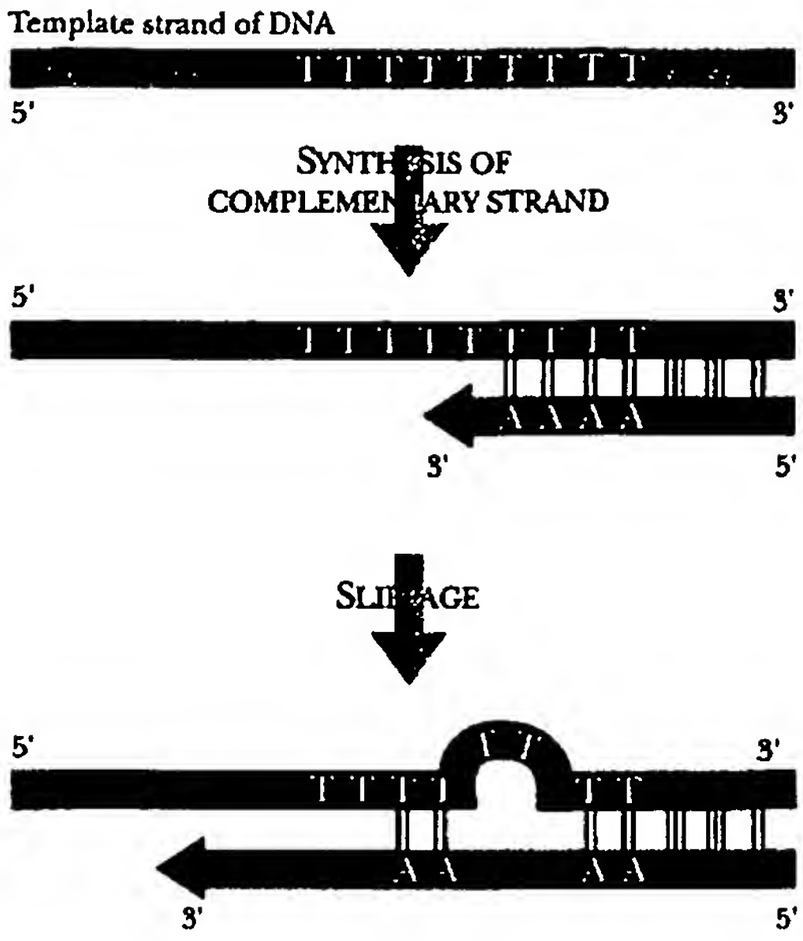


FIGURE 13.17 ***Strand Slippage Creates Small Insertions or Deletions***

The template strand of DNA shown contains numerous thymines (T) in a row (shown in yellow). When replication occurs thymine pairs with adenine (A). However, a long tract of identical bases may cause confusion and some thymines may slip and pair out of register. The extra T residues of the template strand do not pair and form a bulge. In the case shown a small deletion of two bases has occurred.

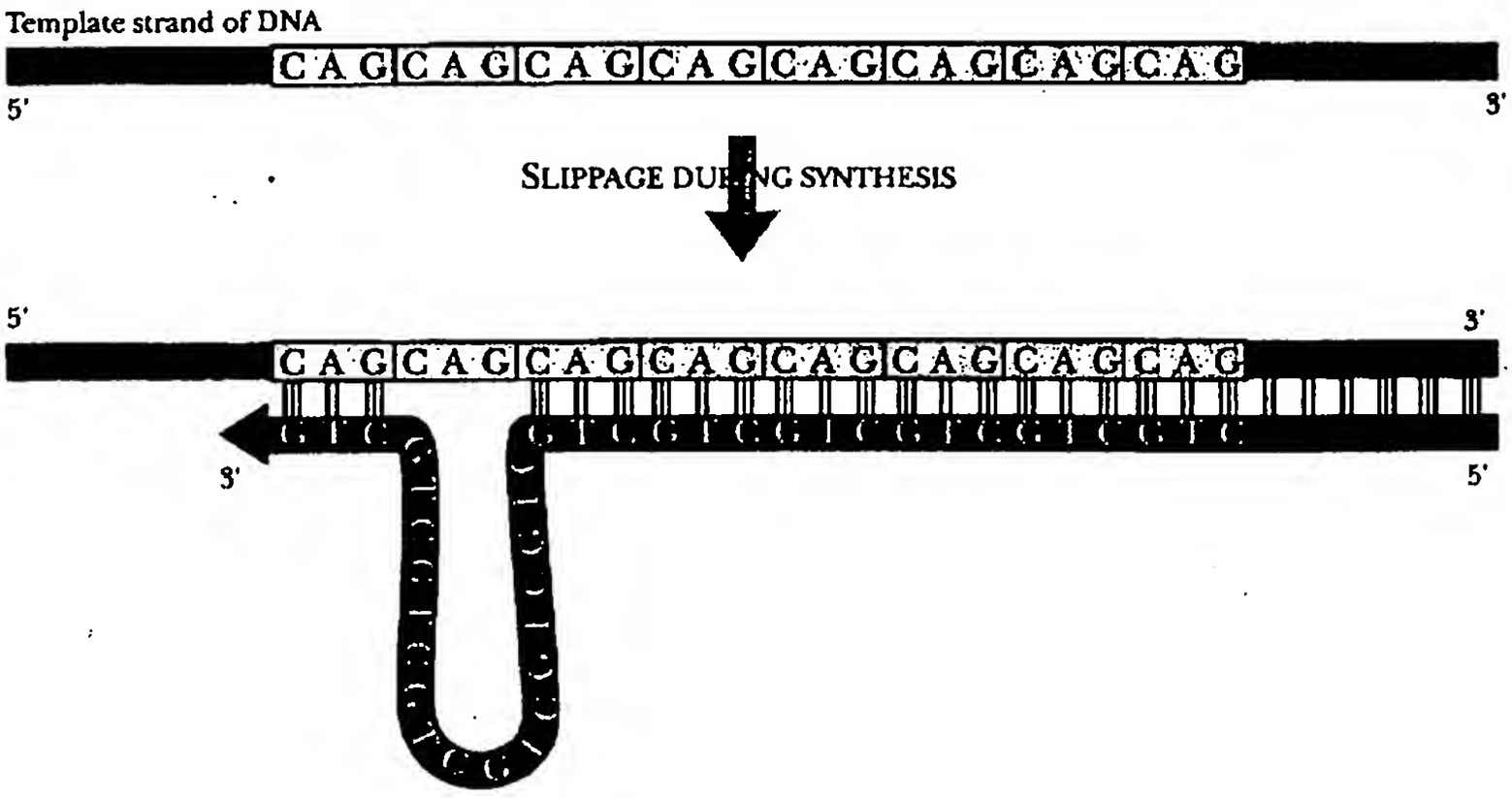


FIGURE 13.18 ***Strand Slippage of Trinucleotide Repeat***

Multiple trinucleotide repeats, such as CAG, may cause strand slippage during DNA replication. In the case illustrated, looping out has occurred in the newly synthesized strand of DNA. The result will be an insertion of six trinucleotide repeats.

become misaligned (Fig. 13.17). Depending on which strand slips, a base may be inserted or omitted during replication.

Slippage may also occur in regions of DNA where there are multiple repeats of a short sequence, perhaps two or three bases (Fig. 13.18). In this case, a whole repeat unit of several bases will be added or deleted. Well known cases occur in the human trinucleotide repeat expansion diseases, such as fragile X syndrome and Huntington's disease. Here copies of a three-base repeat are added or lost due to slippage.

Mutations Can Result from Mispairing and Recombination

Recombination may occurs between closely related sequences of DNA, such as two alleles of the same gene. Many DNA rearrangements, including deletions, inversions, translocations and duplications may result from mistaken pairing of similar sequences followed by recombination. The mechanism of recombination is dealt with in Ch. 14; here, the overall result of mispairing will be considered. If the similar sequences are in the same orientation, mispairing followed by crossing over will generate a duplication on one molecule of DNA and a corresponding deletion on the other (Fig. 13.19).

If two copies of a sequence are on the same DNA molecule but face each other (i.e., are in opposite orientations), mispairing followed by crossing over will generate an inversion (Fig. 13.20). For example, the chromosome of *E. coli* contains seven copies of the genes for ribosomal RNA. Strains of *E. coli* are known in which the whole segment of the bacterial chromosome between two of these rRNA operons has been inverted. Such strains grow slightly slower but nonetheless, are viable.

Spontaneous Mutation Can Be the Result of Tautomerization

However sophisticated DNA polymerase may be, there are chemical limits on the accuracy of DNA replication. Even if DNA polymerase inserts the correct base, errors may still occur. This is due to the **tautomerization** of the bases that constitute DNA. Each base may exist as two possible alternative structures that interconvert. Such structural isomers that exist in dynamic equilibrium are known as tautomers. In each case, one isomer is much more stable and the vast majority of the base is found in this form. However, the less stable alternative tautomer will appear very rarely. If this happens just as the replication fork is passing, the rare tautomer may cause incorrect base pairing.

Thymine has keto and enol tautomers (Fig. 13.21). The common, keto-form pairs with adenine, but the rare enol-tautomer base pairs with guanine. Guanine also has keto and enol tautomers. In this case the rare enol-guanine base pairs with thymine rather than cytosine. Similarly, adenine equilibrates between common amino and rare imino tautomers. The rare imino-adenine base pairs with cytosine instead of thymine. Cytosine alone does not have the potential to introduce mismatches. Although it does have amino and imino tautomers, both pair with adenine. As the temperature increases, the probability that a base is in the incorrect tautomeric state also increases and so, therefore does the mutation frequency.

Spontaneous Mutation Can Be Caused by Inherent Chemical Instability

Although DNA is relatively stable, some of its components do show a low level of spontaneous chemical reaction. Several bases undergo slow but measurable loss of their amino group; i.e., **deamination**. Adenine, guanine and cytosine may all spontaneously deaminate, but by far the most frequent is the deamination of cytosine to give uracil (Fig. 13.22). In addition, the modified base, 5-methyl-cytosine, is especially prone to deamination, so giving "methyl-uracil;" in other words, thymine. The result, in both cases, is the replacement of C by T. Deamination of A (to hypoxanthine) and G (to xanthine) occurs at only 2 to 3 percent of the rate for cytosine. Both hypoxanthine and xanthine usually (but not exclusively) base pair with C, so mutations may be introduced in some cases.

deamination Loss of an amino group
tautomerization Alternation of a molecule, in particular a base of a nucleic acid, between two different isomeric structures

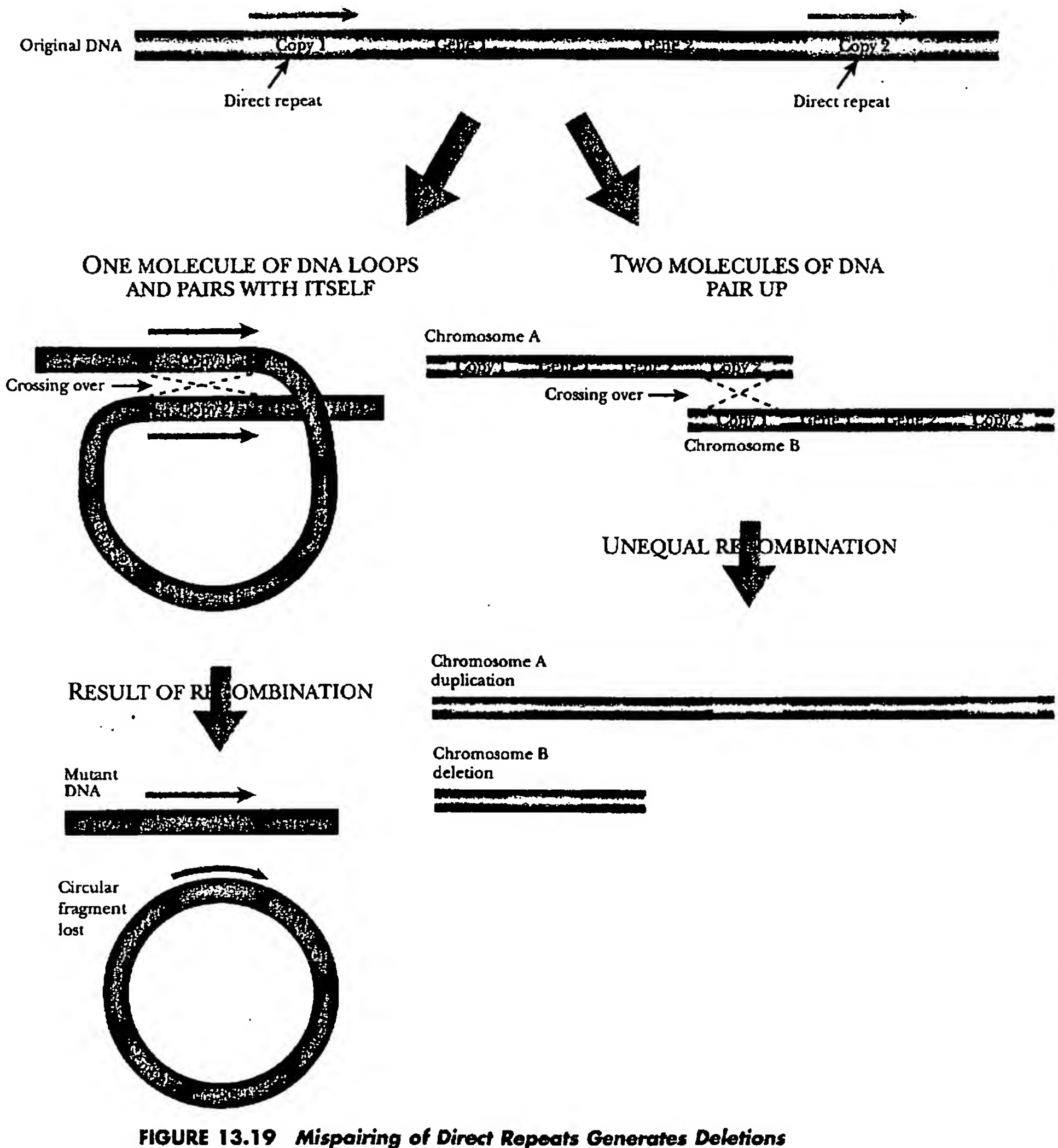


FIGURE 13.19 ***Mispairing of Direct Repeats Generates Deletions and Duplications***

Direct repeats in a DNA molecule may undergo two fates. On the left, the two repeats in a single DNA molecule pair up and recombine. This yields two products; the original DNA molecule suffers a deletion of DNA between the two repeats and a separate circular molecule of DNA is released. On the right, a repeat on chromosome A pairs with another repeat on chromosome B. The result after recombination is a duplication on chromosome A and a deletion of the corresponding region from chromosome B.

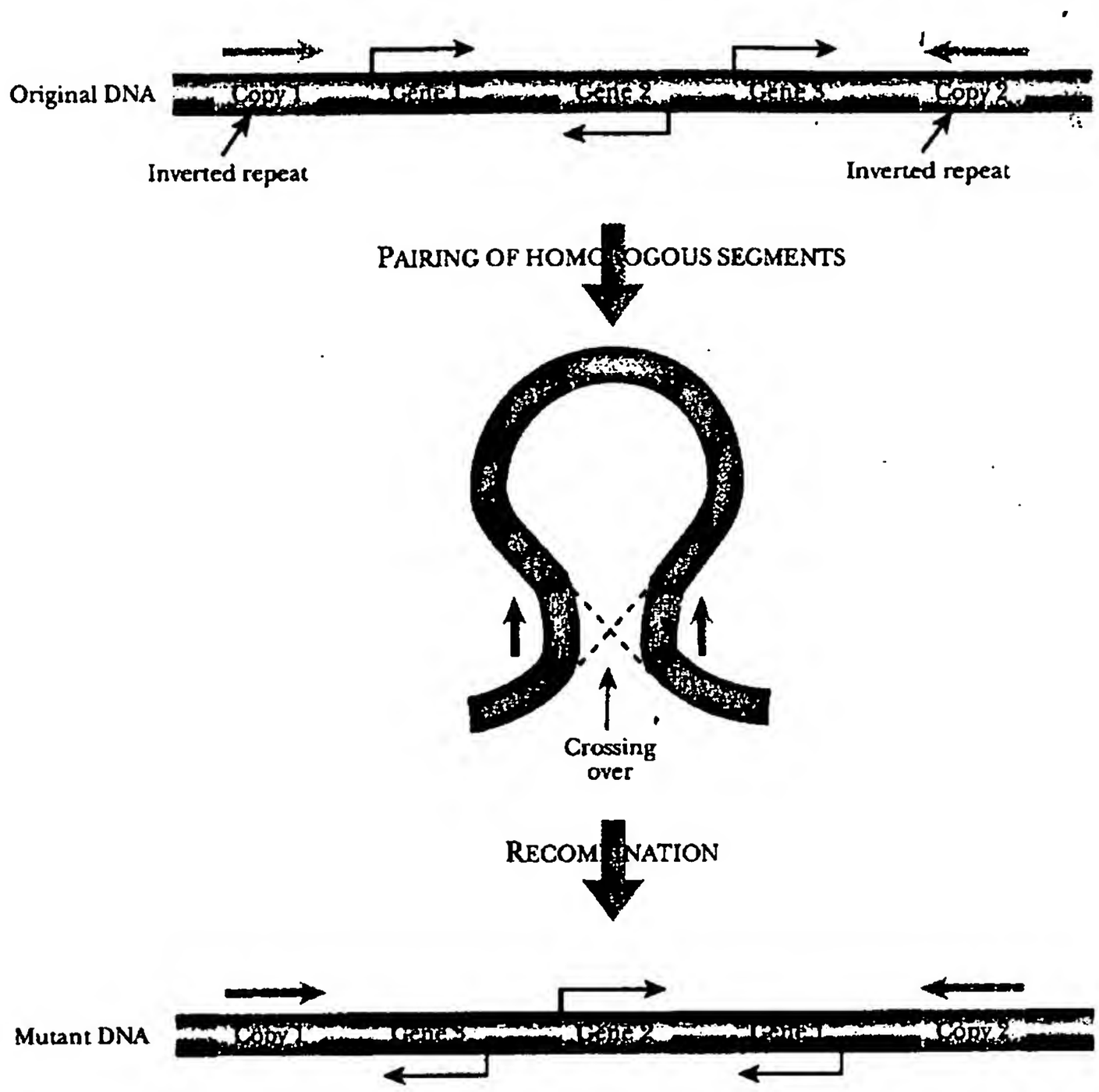


FIGURE 13.20 ***Inversion of DNA by Mispairing of Inverted Repeats***

The DNA molecule shown has two copies of a sequence that are inverted relative to each other. Three intervening genes (Genes 1, 2 and 3) with their directions of transcription (arrows) are also shown. The duplicate sequences may pair up, forming a stem and loop, and undergo recombination. The result is an inversion of the region between the duplicate sequences. This reverses the direction of transcription of the three enclosed genes with respect to the DNA molecule.

Oxidative damage to DNA is also significant. Hydroxyl and superoxide radicals derived from molecular oxygen will attack several bases. The most common target is guanine, which is oxidized to 8-hydroxy-guanine, which pairs preferentially with A. Hence a G · C base pair may be mutated into a T · A pair.

Non-enzymatic methylation of bases occurs at a low frequency. The methyl donor, S-adenosyl-methionine, is normally used by enzymes that attach methyl groups to their substrates. However, it is sufficiently reactive to attack several bases at a low rate spontaneously. The major problem is the formation of 3-methyl-adenine, which tends to block DNA elongation.

Occasionally, the bonds linking the bases of DNA to deoxyribose may spontaneously hydrolyze. This occurs more often with purines than with pyrimidines, generating empty, apurinic sites. Such missing bases tend to block DNA replication and are also an invitation to DNA polymerase to insert an incorrect base.

Mutations Occur More Frequently at Hot Spots

If the same gene is mutated thousands of times, are the mutations all different and are they distributed at random throughout the DNA sequence of that gene? Many of them

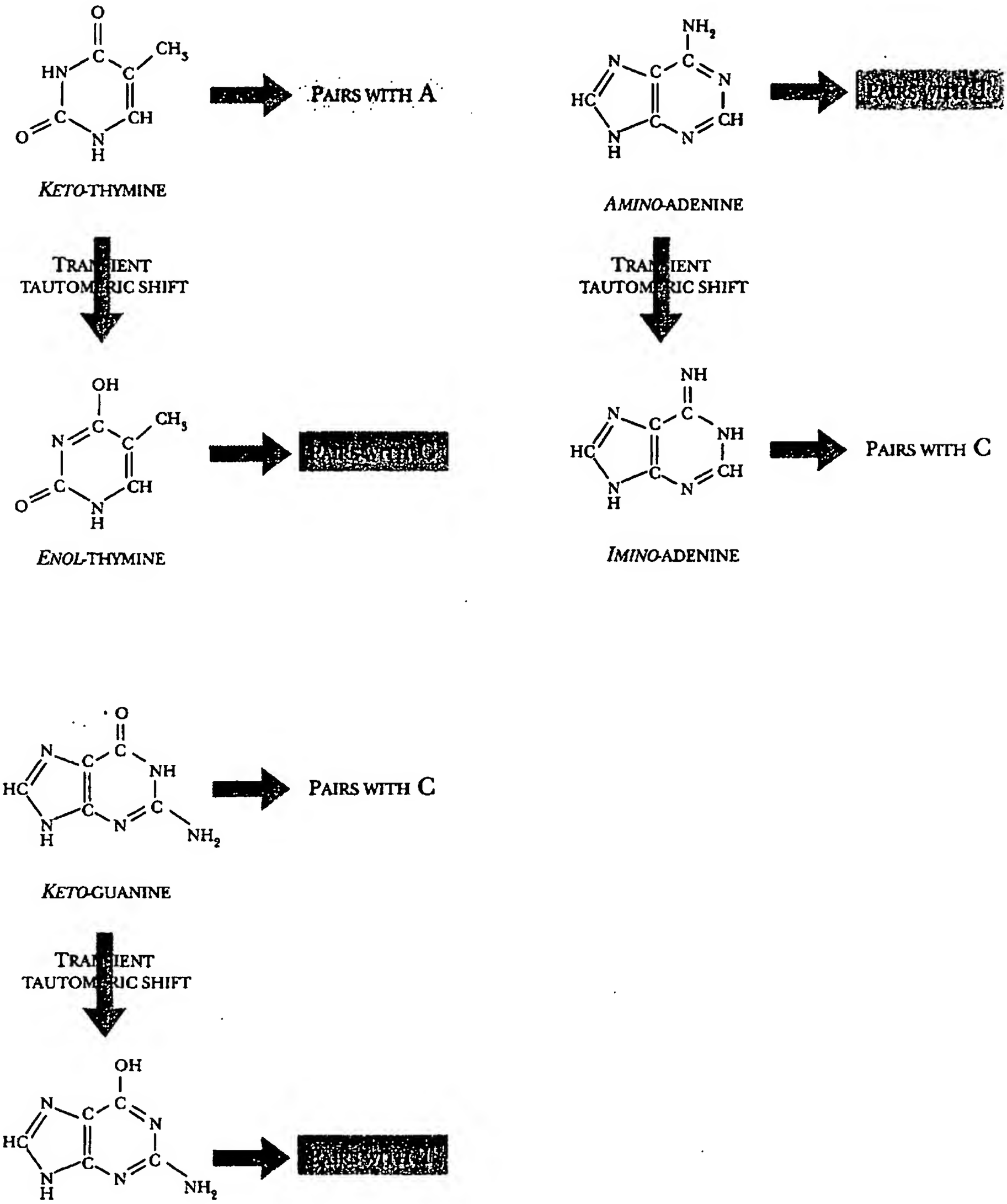
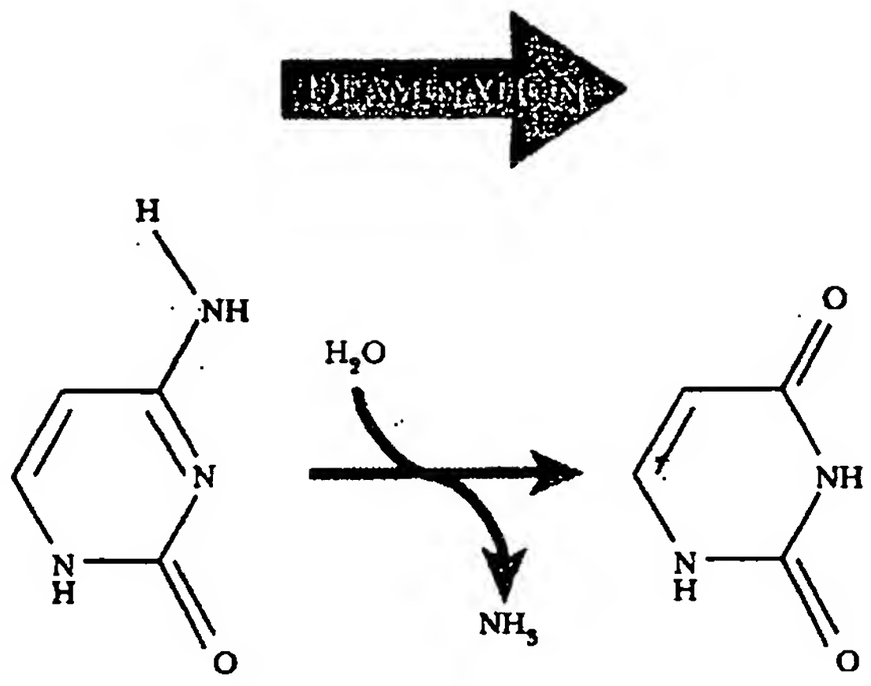


FIGURE 13.21 ***Base Tautomerization May Cause Mismatches***

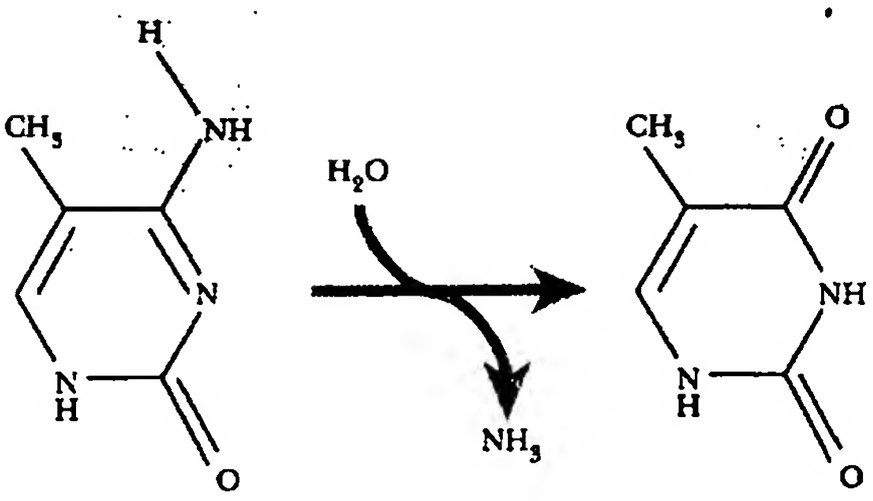
The tautomers of thymine, adenine and guanine are shown. In each instance the lower, short-lived tautomer pairs with the inappropriate base.



CYTOSINE URACIL

FIGURE 13.22
Deamination of Cytosine and 5-Methyl-cytosine

The deamination of cytosine yields uracil and the deamination of the methylated form of cytosine yields thymine—both inappropriate bases.



5-METHYLCYTOSINE THYMINE

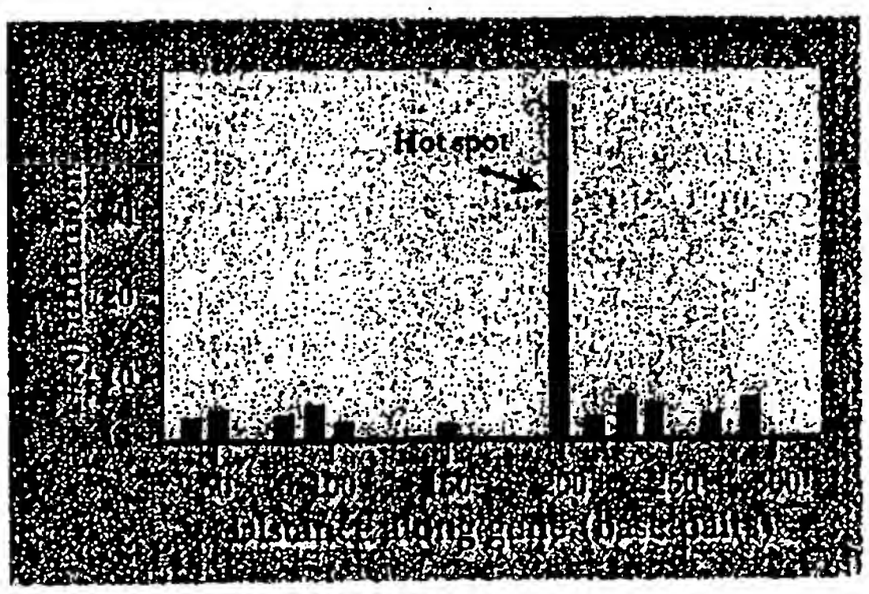


FIGURE 13.23 ***Hot Spots in Distribution of Mutations***

The frequency of mutations along a gene is graphed, showing that one particular location at around 200 base pairs along the gene receives far more mutational events than other regions.

are, but here and there in the DNA sequence are locations where mutations happen many times more often than average (Fig. 13.23). All the mutations occurring at such a site will usually be identical. These sites are called **hot spots**.

Most hot spots are due to the presence of occasional methyl-cytosine bases in the DNA (see Ch. 4 for DNA methylation). These are made from cytosine after DNA synthesis and they pair correctly with guanine, just like normal cytosine. However, every

hot spots Site in DNA or RNA where mutations are unusually frequent

TABLE 13.01 Mutation Rates in DNA Genomes

Organism	Genome Size (kilobases)	Mutation Rate per Generation: Per kb	Per genome (uncorrected)	Per effective genome
Bacteriophage M13	6.4	7.2×10^{-4}	0.005	0.005
Bacteriophage Lambda	49	7.7×10^{-5}	0.004	0.004
Escherichia coli	4,600	5.4×10^{-7}	0.003	0.003
Saccharomyces cerevisiae	12,000	2.2×10^{-7}	0.003	0.003
Caenorhabditis elegans	80,000	2.3×10^{-7}	0.018	0.004
Drosophila	170,000	3.4×10^{-7}	0.058	0.005
Human	3,200,000	5.0×10^{-8}	0.160	0.004

now and then methyl-cytosine spontaneously deaminates to give thymine (=methyl-uracil). This pairs with adenine, not with guanine, and so when the DNA is replicated next, an error results.

Hot spots also occur for deletions, insertions and other major DNA rearrangements. Depending on the mechanism of mutation, certain sequence motifs will favor particular genetic events. As noted above, many rearrangements are due to illegitimate recombination between two nearby regions of DNA with similar sequences.

How Often Do Mutations Occur?

The rates of mutation per generation for several well known organisms are shown in Table 13.01. Defining the rate of mutation is not as obvious as it might appear. Should mutations be expressed as the rate of alterations to the DNA sequence or should they be expressed as the rate of alterations in functional genes? In viruses and bacteria, where most DNA encodes genes, it makes little difference which method of expression is utilized. However, in higher organisms, most DNA is non-coding and alterations in the non-coding DNA rarely have significant effects on the viability of the organism. The concept of "**effective genome**" (i.e., the coding portion of the genome) allows for this and can be used when considering mutation rates. In Table 13.01, the mutation rates cited for the multi-cellular organisms per "effective genome" differ from the raw values. For the single-celled organisms, these two rates are, of course, nearly identical.

All organisms have similar mutation rates, if the effective genome size is considered.

Overall, the more DNA per genome, the lower the mutation rate per kilobase of DNA. Higher organisms have both more accurate DNA replication and more sophisticated DNA repair systems in order to cope with their extra DNA. However, because they have such vast amounts of DNA, the mutation rates per genome are still much higher for the more advanced organisms. In contrast, mutation rates per effective genome are very similar for all organisms, suggesting that this is the level at which evolutionary constraints act. Too high a mutation rate will cause too much damage; too low a rate will fail to provide enough new mutations to drive evolution.

Another ambiguity resides within the term "generation." For single-celled organisms, there is no problem in defining generation, but, again, higher organisms are problematic. Does the term generation refer to an individual cell or of the multicellular organism as a whole? For example, there are many cell generations that transpire between a fertilized egg and the next generation of reproductive cells (i.e., the egg or sperm cells). The values in Table 13.01 refer to each cellular generation.

effective genome The portion of the genome that consists of useful genetic information and ignores the intervening and non-coding DNA. Only applicable to eukaryotic organisms

In the roundworm *Caenorhabditis elegans*, there are only about 10 cell divisions between the zygote and the next generation of germ cells, so the mutation rate per generation of whole animals is 0.04 (about ten times the rate per cell generation). For animals with more cell divisions between generations, this number is much higher—in flies it is 0.13 and for humans 1.6. In flies and humans, the number of cell divisions leading to sperm is significantly greater than the number leading to egg cells. Consequently, sperm carry many more new mutations than eggs and the male parent contributes a greater portion of mutations to the offspring than the mother.

RNA genomes have extremely high mutation rates.

Unlike DNA polymerases, RNA polymerases lack the ability to proofread. Consequently, RNA-based genomes have much higher rates of spontaneous mutation. The mutation rates per genome per generation range from 1 to 5 for small RNA viruses—a rate approximately 1000-fold higher than for the DNA-based cells in Table 13.01. This rate is so high that a significant fraction of the virus particles are defective and it has been estimated that a three-fold higher mutation rate would cause total lack of viability. RNA is only used as the genetic material by certain viruses with relatively small genomes, such as influenza or AIDS. These viruses evade immune surveillance by mutating very rapidly (see Ch. 17). In practice, RNA is not used as the genetic material by any living cell, or even by larger viruses such as bacteriophage T4 or smallpox. Presumably the much higher error rate characteristic of RNA would cause severe problems in organisms where many gene products interact.

Reversions Are Genetic Alterations That Change the Phenotype Back to Wild-type

A reversion that precisely restores the original DNA sequence is highly unlikely.

Obviously, it is possible for DNA that already carries one mutation to be mutated again. There is a small chance that the second mutation will reverse the effect of the first. This process is called **reversion** and refers to the observable outward characteristics of an organism. Reversion is thus a term used to describe a phenotype.

What is the chance a single preselected base will mutate to revert itself to the wild-type? The likelihood that precisely the one base out of millions that were previously mutated will be the very one to mutate again is extremely low. Those rarities where the original base sequence is exactly restored are known as **true revertants**. More often, revertants actually contain a second base change that cancels out the effect of the first one. These are therefore known as **second-site revertants** and the second mutation is known as a **suppressor mutation**.

Most revertants have changes in their DNA that cancel out the effects of the previous mutation.

Not surprisingly, mutations that involve more than a single base change are much less likely to revert. Since part of the original DNA sequence has been completely lost, deletions are completely nonrevertible, at least as far as restoring the original DNA sequence is concerned. Precise reversal of insertions, inversions and translocations is also extremely rare, though not theoretically impossible. Nonetheless, in such cases, reversion is almost always due to compensatory changes in another gene(s).

A second-site reversion can occur if the original mutation was a frameshift mutation caused by the deletion or insertion of a single base. The frameshift mutation alters the reading frame and garbles the protein sequence, as shown in Fig. 13.24. But suppose an extra base is inserted a little way farther along the sequence. This second-site insertion will restore the original reading frame. Although the DNA sequence is not identical to its original state, because of base redundancy the protein has been exactly restored. Similarly, an insertion mutation can be corrected by a second-site deletion.

A less obvious but more frequent case of reversion occurs where the original mutation was a base change. Again, the key to successful reversion is to restore activity to

reversion Alteration of DNA that reverses the effects of a prior mutation
second-site revertant Revertant in which the change in the DNA, which suppresses the effect of the mutation, is at a different site to the original mutation
suppressor mutation A mutation that restores function to a defective gene by suppressing the effect of a previous mutation
true revertant Revertant in which the original base sequence is exactly restored

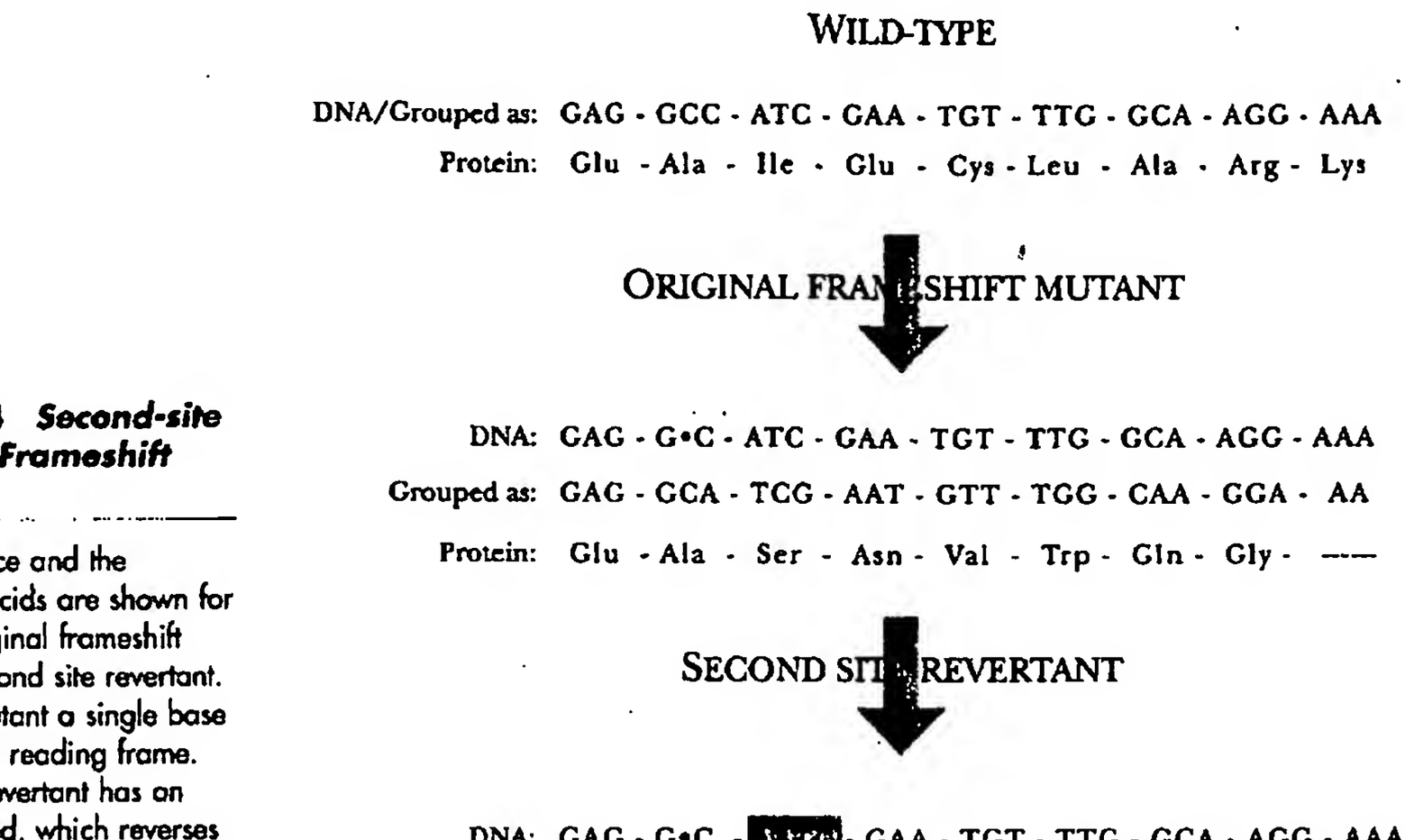


FIGURE 13.24 ***Second-site Reversion of Frameshift Mutation***

The DNA sequence and the encoded amino acids are shown for wild-type, an original frameshift mutant and a second site revertant. In the original mutant a single base deletion alters the reading frame. The second site revertant has an extra base inserted, which reverses the original frameshift. Although the DNA sequence is not identical to the wild type, the amino acid sequence of the protein has been restored.

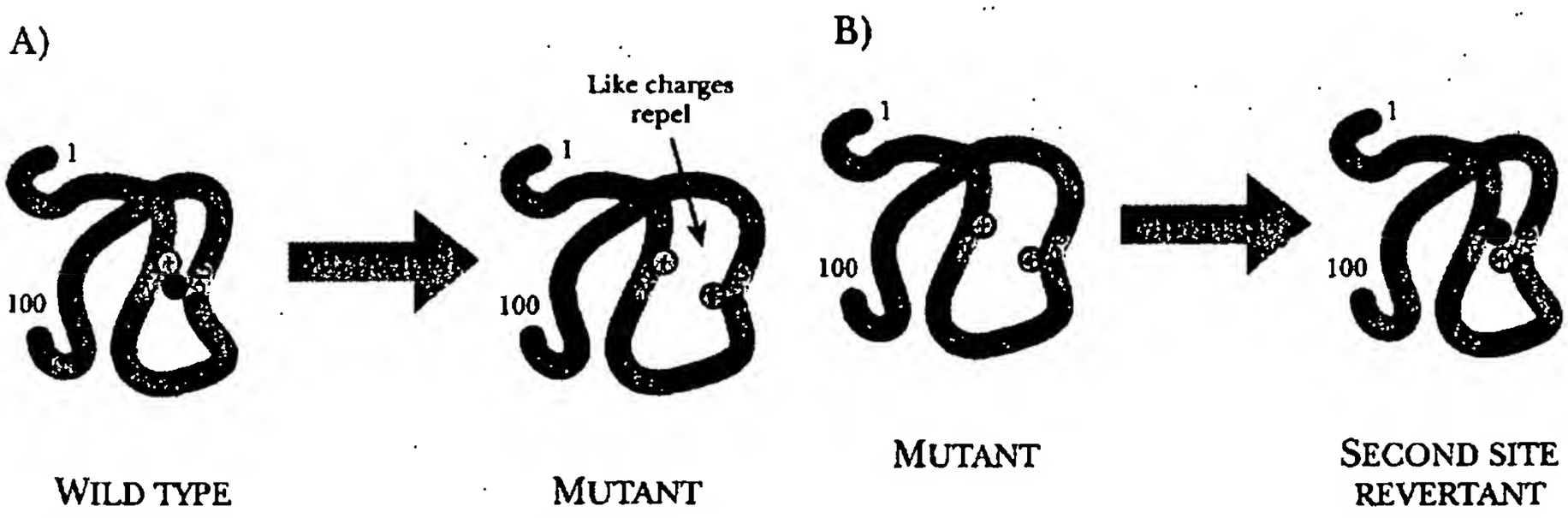


FIGURE 13.25 ***Second-site Reversion of Base Change Mutation***

A) The original mutation alters amino acid #50 from negatively to positively charged. This causes a change in conformation due to charge repulsion. B) A second mutation alters amino acid #25 from positively to negatively charged. This restores the attraction between position #25 and #50 and the protein reverts to its original conformation.

the protein. The precise restoration of the original DNA sequence is less important. Consider a protein whose correct three-dimensional structure depends on the attraction between a positively charged amino acid at, say, position 25 and a negatively charged one at position 50. Suppose the original mutation changes codon 50 from GAA for glutamic acid (negatively charged) to AAA, which encodes lysine, a positively charged amino acid (Fig. 13.25). The folding of the protein is now disrupted due to charge repulsion. A true revertant could be made by replacing AAA with GAA. However, the attraction between residues 25 and 50 can be restored by mutating codon

The stop codons were originally identified by mutations in bacteriophage T4. The first one identified was UAG, the amber codon, which received its name in a curiously convoluted manner. The laboratory of Seymour Benzer at Caltech was looking for a mutation that would allow a certain kind of bacteriophage mutant to grow. Benzer said that whoever identified the mutation, would have it named after him. The mutation was eventually isolated by a student named Harris Bernstein. Since "Bernstein" is German for "amber" UAG was named the amber codon. The second stop codon to be found (UAA) was called "ochre" to keep the color theme. The third stop codon, (UGA) is less common and so the use of "opal" or less often "umber" is less frequent and not fully settled.

25 to give a negatively charged amino acid. This yields a negative charge at position 25 and a positive charge at position 50. Because the attraction between these two regions has been restored, the protein may fold properly again (Fig. 13.25). Will such a revertant protein work correctly? Sometimes yes, sometimes no; it depends on a variety of other factors, such as whether folding is completely restored and whether the alterations damage the active site.

Reversion Can Occur by Compensatory Changes in Other Genes

Sometimes reversion is due to compensatory changes in a completely different gene.

During selection of revertants of a particular gene based on phenotypic differences, a mixture of new mutations will be found. Experimentally this is most easily done using cultures of bacteria or yeast, but the principle applies to higher organisms, too. Occasional true revertants will regain the original DNA sequence. Various possible second site revertants will occur that restore at least some activity to the protein. These may lie within the gene that was originally mutated or in other genes. Restoring activity to a protein by a second mutation within the same gene is sometimes known as **intragenic suppression**. This is to contrast with **extragenic suppression**, where the effects of a mutation are suppressed by a compensatory mutation in a second, quite separate gene.

Extragenic suppression is also called intergenic suppression, to indicate that the two genes involved interact in some way. A whole variety of possible mechanisms exists for such effects. Since an alteration in one gene is making up for a defect in another, extragenic suppression rarely restores function completely.

Two examples of extragenic expression are presented; one due to metabolic compensation and the other due to altered tRNA (see below). Many organisms have multiple genes that code for closely related proteins. For example, the fumarate reductase and the succinate dehydrogenase of bacteria such as *E. coli* both catalyze the same reaction, the interconversion of fumarate with succinate. The fumarate reductase (FRD) is used to make succinate from fumarate during anaerobic growth, whereas the succinate dehydrogenase (SDH) functions in the Krebs Cycle to convert succinate to fumarate during aerobic growth. Although there are slight differences in the DNA and protein sequence, the major difference is in the mode of regulation; the *frd* genes are normally expressed only anaerobically and the *sdh* genes only in the presence of oxygen. Consequently, a mutation in *frd* may be suppressed by a regulatory mutation that allows expression of succinate dehydrogenase anaerobically. Conversely, a mutation in *sdh* may be suppressed by a regulatory mutation that turns on fumarate reductase in the presence of oxygen.

extragenic suppression Reversion of a mutation by a second change that is within another distinct gene
intragenic suppression Reversion of a mutation by a second change at a different site but within the same gene

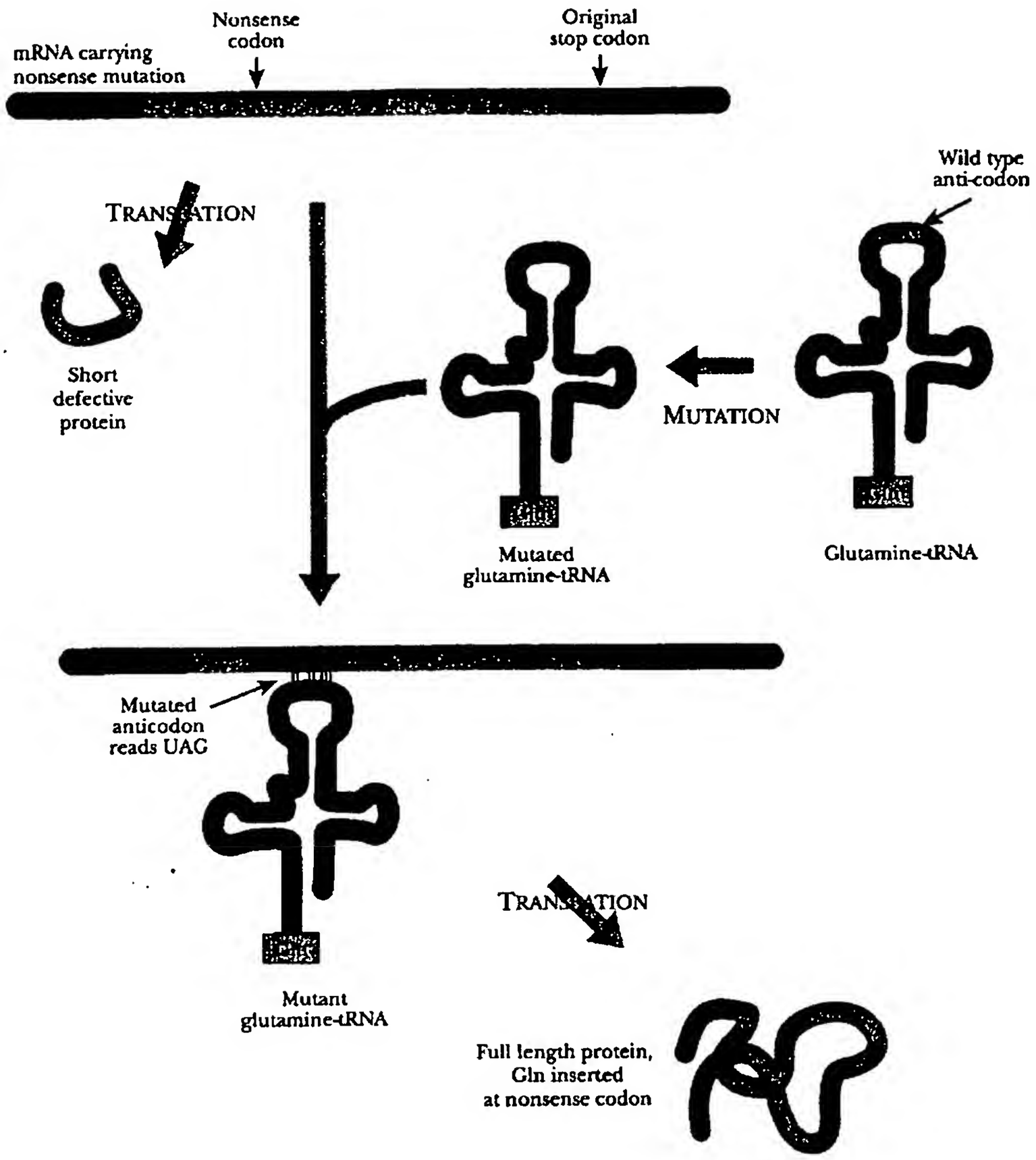


FIGURE 13.26 ***Mechanism of Nonsense Suppression***

A gene containing a nonsense codon suffers a premature stop during translation and a short defective protein is made. However, a tRNA whose anticodon is mutated (from GUC to AUC) can recognize the stop codon and insert an amino acid (glutamine in this case). A full length protein will be made that has only one amino acid different from the original wild type.

Altered Decoding by Transfer RNA May Cause Suppression

Mutant tRNA molecules are known that read stop codons and insert amino acids. This suppresses nonsense mutations.

Nonsense mutations can be suppressed by alterations in tRNA. As discussed above, a nonsense mutation occurs when a codon that should code for an amino acid is changed to a stop codon. This results in a truncated and usually nonfunctional protein. Such a defect may be suppressed, at least partially, by changing the anticodon sequence of a tRNA molecule so that it recognizes the stop codon instead. Consider the stop codon UAG. Altering the anticodon of $tRNA^{Gln}$ from GUC (reading CAG for Gln) to AUC will make it recognize UAG instead. Such an altered tRNA will insert glutamine wherever it finds a UAG stop codon (Fig. 13.26).

Such altered tRNA molecules are known as **suppressor tRNAs**. The UAG stop codon is known as *amber* and the UAA stop codon as *ochre*. The UGA stop codon has no universally accepted name, but is sometimes called *opal* or *umber*. Amber suppressors are mutant tRNAs that read UAG instead of their original codon. Ochre suppressor tRNAs read both UAA and UAG due to wobble. Opal suppressors are rare.

Suppressor tRNA mutations can only occur if a cell has more than one tRNA that reads a particular codon. One may be mutated while the other must carry out the original function; otherwise, the loss of the original tRNA would be lethal. In practice, cells often have multiple tRNA genes and so suppressor mutations are reasonably common, at least in microorganisms. Bacterial suppressor mutations have been found in tRNAs for glutamine, leucine, serine, tyrosine and tryptophan. The amino acid inserted by the suppressor tRNA may be identical to the original amino acid whose codon mutated to give the stop codon. In this case, the protein made will be fully restored. Alternatively, a different amino acid may be inserted and a partially active protein may be produced.

Remember that stop codons are normally recognized by release factor, and have no cognate tRNAs. Since suppressor tRNA competes with release factor, suppression is never complete and typically ranges from 10 to 40 percent. This may provide enough of the suppressed protein for the cells to survive. However, the suppressor tRNA will also suppress other stop codons in the same cell and so generate longer (and incorrect) versions of many proteins whose genes were never mutated. Not surprisingly, cells with suppressor mutations grow more slowly. Only bacteria and lower eukaryotes (e.g., yeasts, roundworms) can tolerate suppressor mutations. In both insects and mammals, suppressor mutations are lethal.

Frameshift suppressor tRNAs are also occasionally found among bacteria. These mutant tRNA molecules have an enlarged anticodon loop and a four-base anticodon. This enables them to insert a single amino acid by reading four bases in the mRNA. They can suppress the effects of a frameshift mutation that was caused by the insertion of a single extra base. Frameshift suppressor tRNAs with five-base anticodons have been made artificially, but have not been isolated naturally.

Mutagenic Chemicals Can Be Detected by Reversion

Screening for reversion using well characterized mutations allows the detection of mutagenic chemicals.

Testing chemicals for tumor formation in animals is expensive and time-consuming. However, since cancer is due to DNA alterations, most carcinogens are in fact also mutagens. Consequently, chemicals suspected of being carcinogenic are routinely screened for possible mutagenic effects by testing against bacteria. The **Ames test** makes use of multiple strains of the bacterium *Salmonella typhimurium* carrying well characterized mutations in the genes for histidine synthesis. It is used routinely by industry and government agencies to screen food colorings and preservatives, cosmetics such as hair dyes, and many other industrial chemicals for possible mutagenic effects.

Mutants of *Salmonella typhimurium* carrying mutations in the *his* genes can no longer make the amino acid histidine and cannot grow unless given histidine. When large numbers of these mutant bacteria are placed on growth medium lacking histidine, just a handful of colonies appear. These are revertants, and since reversions are merely mutations back to the original state, the frequency of reversion is also increased by mutagenic agents. To test a suspect chemical, samples of *Salmonella his* mutants are mixed with the agent and then plated onto minimal medium with just a trace of histidine. The amount of histidine added is growth-limiting. Therefore the bacteria can only divide a few times and run out of histidine before making visible colonies. If the added

Ames test Test for mutagenic activity that makes use of bacteria
suppressor tRNA A mutant tRNA that recognizes a stop codon and can insert an amino acid when it reads a stop codon on the mRNA

chemical does induce mutation then His+ revertants will be formed during these few cell divisions, then each resulting revertant can grow into a visible colony. Different types of original mutations, for example, base changes or frameshift mutations, are used to screen for different classes of mutagenic agents.

The *Salmonella typhimurium* strains used in practice for mutagen testing have several alterations that make them more sensitive to mutagens. Firstly, they carry mutations that make the bacterial outer membrane more permeable to large and/or hydrophobic molecules. Secondly a variety of alterations have been made to inactivate bacterial DNA repair mechanisms (see Ch. 14). For example, the *uvrB* gene may be deleted to eliminate excision repair of DNA.

Certain chemicals (pro-mutagens) are only mutagenic after metabolic conversion to active derivatives. In animals this is usually due to liver enzymes such as cytochrome P450 that are intended to detoxify harmful chemicals by oxidation. When testing for pro-mutagens, an extract containing such rat liver enzymes is mixed with the bacteria in the Ames test. Recently, genes for some variants of human cytochrome P450 have been cloned and successfully expressed in the *Salmonella* strains used for mutagen testing. The resulting bacteria synthesize the liver enzymes internally and are much more sensitive in their response to pro-mutagens.

Experimental Isolation of Mutations

The emergence of molecular biology has relied greatly on the analysis of mutations in a variety of organisms. Obviously, such analysis requires a supply of mutants to work with. These may be obtained by a wide range of approaches. Mutations may be spontaneous or artificially induced. Artificial mutagenesis may be carried out using living organisms (*in vivo* mutagenesis) or performed using isolated DNA (*in vitro* mutagenesis). In addition, some means is needed for identifying mutants, whether they occur spontaneously or are made deliberately.

Since the frequency of spontaneous mutations is low, it is only possible to rely on this source of mutations when a population of millions of organisms can be surveyed in a reasonable time. Consequently, this approach is largely restricted to bacteria and single-celled eukaryotes, such as yeast. For a gene of average size (~1000 bp) the mutation rate is approximately 0.5 per million per generation in bacteria such as *E. coli* (see Table 13.01). Thus, a typical culture of several 1,000 million bacteria per ml that has resulted from several generations of growth may contain half a dozen spontaneous mutants per million cells (or several thousand mutants per ml of culture) affecting any particular gene of interest (assuming such mutations are not lethal).

From a practical viewpoint, the problem is not so much causing mutations as finding and isolating the ones desired.

The problem then becomes how to isolate these mutants. It is clearly impractical to examine millions of microorganisms individually. Therefore the isolation of spontaneous mutants relies on some form of direct selection. Usually, samples of the culture are spread on the surface of solid medium designed to allow only the desired mutants to grow. For example, mutations in DNA gyrase make bacterial cells resistant to quinolone antibiotics, such as nalidixic acid. Therefore medium containing nalidixic acid kills the vast majority of bacteria and can be used to isolate gyrase mutants. Sometimes bacterial cultures may be enriched for the required mutation by growth for several generations in liquid selective medium before transferring to solid selective medium for the final isolation. This increases the proportion of the required mutants in the population.

A large number of direct selections have been used to isolate bacterial mutants. The major categories of selection and the kinds of genes affected are as follows:

A. Resistance to antibiotics. Alterations in genes whose products are targets of the antibiotic or that are involved in entry of the antibiotic into the cell. For example, streptomycin resistance selects alterations in ribosomal protein S12, rifampicin resistance selects alterations in RNA polymerase, nalidixic acid resistance selects mutations in DNA gyrase.

B. Resistance to analogs of metabolites. Alterations in genes whose products are involved in synthesis, degradation or transport of the metabolite. For example, chlorate (an analog of nitrate) selects mutants defective in nitrate reductase, chloroethanol selects mutants defective in alcohol dehydrogenase, various selenium compounds select mutants with altered sulfur metabolism.

C. Resistance to bacteriophage. Alterations in genes encoding bacteriophage receptor or components needed for entry of viral DNA. For example, resistance to bacteriophage lambda selects for loss of LamB protein on cell surface, resistance to bacteriophage T1 selects for loss of TonB protein needed to energize viral DNA entry.

D. Growth in the absence of certain metabolic supplements. Usually used to select revertants from mutants defective in synthesis of amino acids, nucleotides, vitamins etc.

E. Growth on certain substances as carbon source. Usually used to select revertants from mutants defective in metabolism of sugars, organic acids by known pathways. Growth on novel compounds is sometimes selected. For example, selection for growth of *E. coli* on propanediol selects for aerobic expression of a pathway normally only expressed anaerobically during the fermentation of deoxysugars.

In Vivo versus *In Vitro* Mutagenesis

The frequency of mutation can be greatly increased by a variety of chemical agents or certain types of radiation, as already described in the earlier part of this chapter. These agents are often used deliberately on living cells (*in vivo* mutagenesis). In particular, sublethal concentrations of mutagenic chemicals may be added directly to growing cultures of bacteria or single celled eukaryotes. It is also possible to treat higher organisms, though the approach is limited by the necessity of screening large numbers of large organisms. Nonetheless, relatively small multi-cellular organisms, such as flies and roundworms, have been successfully mutagenised by this approach.

Alternatively, purified DNA may be treated with the mutagenic agent (*in vitro* mutagenesis—see below). In this case, the DNA must be transformed back into the bacteria before screening for mutations is performed.

Mutations in certain genes (mutator genes) involved in DNA synthesis and repair can themselves increase the frequency of mutation. The presence of such defects may be used to generate mutations in bacteria and some single celled eukaryotes. However, such defects tend to be lethal in more complex organisms.

Insertion and deletion mutations may also be made *in vivo* by using transposable elements. Genetic elements such as insertion elements, transposons or bacteriophage Mu insert semi-randomly into host DNA. If the insertion point lies within a host cell gene, the result is an insertion mutation that inactivates the gene in question. Certain transposable elements excise themselves from DNA inaccurately, leaving behind deletions at the point of prior insertion. The behavior of transposable elements is described in more detail in Ch 15, Mobile DNA.

If the frequency of mutation has been increased to where several bacteria per thousand carry mutations in the gene of interest, it is possible to identify mutants by screening (as opposed to selection). A culture of bacteria that has been mutagenized is diluted and samples are spread onto the surface of solid media so as to give single colonies (between 50 and a few hundred colonies per plate are typically obtained). The colonies are then screened by a variety of techniques such as indicator media or replica plating. The term "phenotypic screening" refers to the analysis of the mutants by their phenotype, rather than by examination of the genes.

Bacterial mutants are often screened for their responses on a variety of specialized culture media.

Indicator media are growth media that change color in response to metabolic reactions. The simplest show alterations in pH due inclusion of a pH indicator in the medium. These are typically used to monitors sugar breakdown by bacteria as this generates acid and so alters the pH of the growth medium. Redox indicators, such as

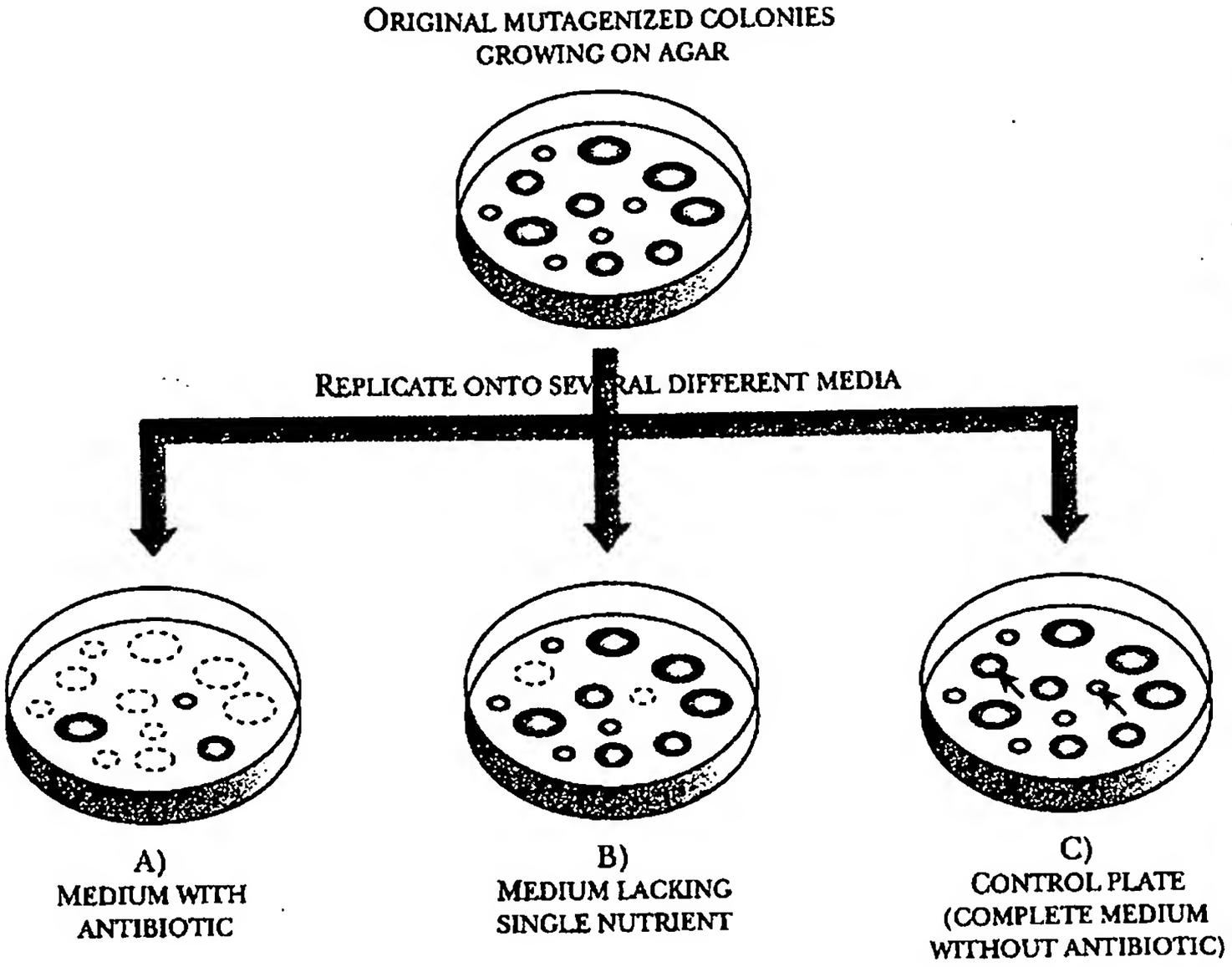


FIGURE 13.27 *Replica Plating*

After treatment with mutagens, bacterial colonies are grown on normal medium in a Petri dish. Replicas are made using pads of velvet or filter paper to transfer samples from each original colony simultaneously to fresh media. Growth on the different media is compared. Colonies present only on a medium with a positive selection, such as an antibiotic, may be retrieved directly (A). However, colonies missing from a particular test medium, such as medium lacking a particular nutrient (B) must be retrieved from the control plate (C). Since these mutants will look identical to the other colonies, they are identified by their position, as indicated by the arrows.

tetrazolium dyes, respond to oxidation reduction reactions. They may be used to monitor the ability of bacteria to oxidize certain growth substrates.

More specific indicators, such as substrates for individual enzymes, may also be used. For example, X-gal is a substrate for β-galactosidase that releases a blue dye when split by this enzyme (see Ch. 7). Bacterial colonies expressing significant levels of β-galactosidase turn blue in the presence of X-gal. One of the most cheerful indicators is the use of selenium salts. Bacteria that can reduce selenate or selenite accumulate granules of elemental selenium that are bright red. In this case several steps of a specific pathway, rather than just one enzyme, are involved. Note that for an indicator system to work well, the colored product must be insoluble, otherwise it will diffuse through the agar and the color will no longer be localized to the colony that performed the reaction.

Replica plating allows screening for mutants that fail to grow under the test conditions.

Replica plating is another widely used form of phenotypic screening. It is particularly useful when searching for mutants that have lost the ability to grow under certain conditions. The same mutagenized bacterial colonies are tested for growth on a variety of media and colonies that fail to grow on the medium of interest are kept. For example, media with different carbon sources, growth supplements or growth inhibitors may be used. Since it is not possible to subculture a colony that failed to grow on a test medium, the mutagenized colonies are first grown on normal (i.e. non-selective) medium. Filter paper or velvet pads are then pressed onto the colonies and pick up bacteria from each original colony. Bacteria corresponding to each colony are then transferred to assorted test media by pressing the filter paper or velvet pads onto the surface of the fresh medium (Fig. 13.27). This technique preserves the arrangement of the colonies on the agar and allows colonies missing on the test medium to be retrieved from the original master plates.

Site-Directed Mutagenesis

DNA may be manipulated in the test tube and the altered DNA construct may then be inserted into the target organism. The simplest form of such *in vitro* mutagenesis is

TABLE 13.02 Techniques used for *In Vitro* Mutagenesis

Chemical mutagenesis of cloned DNA
The gene to be mutagenized is cloned onto a suitable vector, usually a plasmid. DNA carrying the target gene is extracted and purified and treated with a chemical mutagen *in vitro*. The altered DNA is then transformed back into the original organism and screening is carried out to identify organisms that received a mutant version of the gene.

Gene disruption by restriction and ligation (See Ch. 22, Recombinant DNA)
A DNA cassette, often carrying a gene for resistance to some antibiotic to allow selection, is inserted into the target gene by using restriction enzymes and DNA ligase. This approach is often used if convenient restriction sites are available. If not, then PCR-based introduction of extra DNA is a good alternative.

In vitro DNA synthesis (See Ch. 24, Genomics)
Single stranded DNA is sometimes generated for sequencing by using M13 vectors. In vitro DNA synthesis may be performed using such ssDNA as template using T7 polymerase and a supply of nucleoside triphosphates. DNA polymerization may be initiated using artificially synthesized primers whose sequence has been altered by a few bases. This will generate a mutagenized product that incorporates these changes. This technique has largely been replaced by PCR based methods.

PCR based techniques (See Ch. 23, PCR)

a) Introduction of Specific Base Changes
Using PCR primers whose sequence has been altered will generate a PCR product that incorporates these changes.

b) Localized random mutagenesis
Manganese ions cause errors in PCR reactions. Hence random mutations may be introduced into the segment of DNA being amplified.

c) Generation of Insertion or Deletion by PCR
Using PCR primers that include sequences homologous to the target location allows replacement of a region of chromosome with a segment of DNA generated by PCR.

Transgenic technology
Transgenic technology creates genetically modified organisms. It may therefore be regarded as a form of mutagenesis. Extra DNA sequences may be introduced from other organisms, by a variety of techniques.

DNA alterations are often constructed by a variety of genetic engineering techniques.

to treat purified DNA with mutagenic chemicals or radiation. However, a variety of more sophisticated techniques have been used to deliberately construct mutations utilizing genetic engineering technology. These techniques are usually known as **directed mutagenesis** (or, sometimes, **site-directed mutagenesis**, when the site of mutation is carefully controlled). Some techniques introduce changes in one or a few bases, whereas others involve more drastic alterations. Some in vitro techniques generate semi-random base changes whereas others are extremely specific. Assorted methods have been used, in particular PCR (see Ch. 23) is now widely used and has replaced many of the older approaches. These applications are discussed together with the appropriate techniques later in this book and are summarized here for reference in Table 13.02.

directed mutagenesis Deliberate alteration of the DNA sequence of a gene by any of a variety of artificial techniques
site-directed mutagenesis Deliberate alteration of a specific DNA sequence by any artificial technique